

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:39 ; Search time 2585.24 Seconds
(without alignments)
9747.064 Million cell updates/sec

Title: US-09-788-074-4
Perfect score: 662
Sequence: 1 accgcgcagcagaccctc.....taataaatgtgttctgtcct 662

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	658.4	99.5	692	4	BM670233 UI-E-DWI-
2	657	99.2	712	5	BU533478 AGENCOURT
3	655.4	99.0	662	5	BX385751 BX385751
4	655.2	99.0	717	4	BG719235 602690296
5	650.8	98.3	820	4	BG678687 602624485
6	647	97.7	786	1	AV727829 AV727829
7	644.2	97.3	708	4	BG715536 602675842
8	643.8	97.3	748	5	BU943612 AGENCOURT
9	643	97.1	666	4	BM789868 K-EST0069
10	640.8	96.8	843	5	BU601920 AGENCOURT
11	640.4	96.7	656	1	AV724434 AV724434
12	639.4	96.6	686	6	CB530038 UI-H-FT2-
13	636	96.1	655	5	BQ446473 UI-H-EUI-
14	632.8	95.6	654	6	CD370550 UI-H-FT1-
15	632.4	95.5	799	5	BU532649 AGENCOURT
16	632.2	95.5	668	1	AV726696 AV726696
17	631.4	95.4	664	4	BG179048 602330759
18	628.2	94.9	873	5	BU941286 AGENCOURT
19	627	94.7	631	5	BX404021 BX404021
20	626.8	94.7	644	1	AI080481 0871d05.x
21	625	94.4	668	4	BG168846 602320168
22	624.8	94.4	769	5	BU955447 AGENCOURT
23	624.4	94.3	684	4	BF981201 602310488
24	620	93.7	642	5	BU783805 in09g12.x

25	620	93.7	664	2	BF341608	BF341608 602016030
26	619.6	93.6	761	5	BU963854	BU963854 AGENCOURT
C 27	618.8	93.5	654	4	BM665113	BM665113 UI-E-CQ1-
28	618	93.4	632	6	CB124107	CB124107 K-EST0172
29	615.6	93.0	781	5	BU953727	BU953727 AGENCOURT
30	615.4	93.0	629	4	BM695640	BM695640 UI-E-CQ1-
31	614	92.7	625	4	BM695742	BM695742 UI-E-CQ1-
32	613.4	92.7	615	6	CD679571	CD679571 hq13b11.y
33	612.8	92.6	653	4	BM459183	BM459183 AGENCOURT
34	611	92.3	634	6	CD677981	CD677981 ho25c09.y
C 35	610	92.1	646	4	BM666524	BM666524 UI-E-CQ1-
C 36	608.6	91.9	644	1	AI857611	AI857611 wk95d09.x
37	606	91.5	632	4	BM791091	BM791091 K-EST0071
38	605	91.4	653	2	BF035373	BF035373 601458587
39	603.2	91.1	678	7	CN644162	CN644162 ILLUMIGEN
C 40	602.8	91.1	659	1	AI338767	AI338767 qt53d02.x
41	601.6	90.9	666	5	BX338755	BX338755 BX338755
42	600	90.6	747	5	BU955273	BU955273 AGENCOURT
43	598.4	90.4	614	6	CB138668	CB138668 K-EST0191
C 44	598.2	90.4	654	2	AW009545	AW009545 w884d08.x
45	597.6	90.3	653	6	CD389254	CD389254 AGENCOURT

ALIGNMENTS

RESULT 1
BM670233/c
LOCUS BM670233 692 bp mRNA linear EST 27-FEB-2002
DEFINITION UI-E-DWI-ahc-h-08-0-UI.e1 UI-E-DWI Homo sapiens cDNA clone
ACCESSION BM670233
VERSION BM670233.1 GI:18980130
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 692)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bonaldo, M.P., Lennon, G. and Soares, M.B.
Normalizaion and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=yes.
Location/Qualifiers
1. .692
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DWI-ahc-h-08-0-UI"
/tissue_type="adult"
/dev stage="adult"
/lab host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site:1: EcoR I; Site:2: Not I;
UI-E-DWI is a normalized cDNA library containing the


```

Db      318  |||||TGCTGCACTTTACAGATGGTGCATTTGTTTACGATCAGGATGGGATGGGAGGAAG 377
Qy      363  CACAATAACAGAAAAATTGAAGATGGGAAATAGTGGTGGAGTGTGTCAATGAACAATGT 422
Db      378  CACAATAACAGAAAAATTGAAGATGGGAAATAGTGGTGGAGTGTGTCAATGAACAATGT 437
Qy      423  CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 482
Db      438  CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 497
Qy      483  GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCCCATCTCTTTCTTCT 542
Db      498  GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCCCATCTCTTTCTTCT 557
Qy      543  TTTTCTTTTTCATTTTCAATTTATCTTTATCATATAAACAATTTTACATGACGCTATTT 602
Db      558  TTTTCTTTTTCATTTTCAATTTATCTTTATCATATAAACAATTTTACATGACGCTATTT 617
Qy      603  CAAAGTGTGTGGAATTAATGAGATCATCCCTTGGTTGTTAATAAATGTTTGT 659
Db      618  CAAAGTGTGTGGAATTAATGAGATCATCCCTTGGTTGTTAATAAATGTTTGT 674

```

```

RESULT 3
BX385751
LOCUS      662 bp mRNA linear EST 28-APR-2004
DEFINITION Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
            cDNA clone CS0DC028YA01 5-PRIME, mRNA sequence.
ACCESSION  BX385751
VERSION     BX385751.2 GI:46831903
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 662)
REFERENCE   Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30457359.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            Location/Qualifiers
              1. .662
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DC028YA01"
                /tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
                /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.
                Library was normalized."

```

```

FEATURES
            source
            Query Match      99.0%; Score 655.4; DB 5; Length 662;
            Best Local Similarity 99.4%; Pred. No. 1.8e-148;
            Matches 656; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy      3  CGCGCAGCAGACCCCTCTCTGACGCCCGCCGCCGCCACCCACCATGCCACAGTTCA 62
Db      1  CGCGCAGCAGACCCCTCTCTGACGCCCGCCGCCGCCACCCACCATGCCACAGTTCA 60
Qy      63  GCAGCTGGAAGGAGATGGCGCTGGTGACAGCAAGGCTTTGATGAATACATGAAGGA 122

```

```

Db      61  GCAGCTGGAAGGAGATGGCGCTGGTGACAGCAAGGCTTTGATGAATACATGAAGGA 120
Qy      123  GCTAGAGTGGGAATAGCTTTTGGAAAAATGGGCGCAATGGGCAAGCCAGATTTGTATCAT 182
Db      121  GCTAGAGTGGGAATAGCTTTTGGAAAAATGGGCGCAATGGGCAAGCCAGATTTGTATCAT 180
Qy      183  CACTTGTGTATGTTAAACCTCACCATTAATAAATGAGAGCAGCTTTGAAAAACAACAGTT 242
Db      181  CACTTGTGTATGTTAAACCTCACCATTAATAAATGAGAGCAGCTTTGAAAAACAACAGTT 240
Qy      243  TTCTTTTACCCCTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAAAATCTCAGAC 302
Db      241  TTCTTTTACCCCTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAAAATCTCAGAC 300
Qy      303  TGTCTGCACTTTTACAGATGGTGCATTTGTTTTCAGCATCAGGAGTGGGATGGGAAGGAAAG 362
Db      301  TGTCTGCACTTTTACAGATGGTGCATTTGTTTTCAGCATCAGGAGTGGGATGGGAAGGAAAG 360
Qy      363  CACATAACAGAAAAATTAAGAGATGGGAAATTTAGTGGTGGAGTGTGTCAATGAACAATGT 422
Db      361  CACAATAACAGAAAAATTAAGAGATGGGAAATTTAGTGGTGGAGTGTGTCAATGAACAATGT 420
Qy      423  CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 482
Db      421  CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 480
Qy      483  GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTCTTTCTTCT 542
Db      481  GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTCTTTCTTCT 540
Qy      543  TTTTCTTTTTCATCTCTGTTCAATTTATCTTTATCATATAAACAATTTTACATGACGCTATTT 602
Db      541  TTTTCTTTTTCATCTCTGTTCAATTTATCTTTATCATATAAACAATTTTACATGACGCTATTT 600
Qy      603  CAAAGTGTGTGGAATTAATGAGATCATCCCTTGGTTAATAAATAAATGTTTGTGCT 662
Db      601  CAAAGTGTGTGGAATTAATGAGATCATCCCTTGGTTAATAAATAAATGTTTGTGCT 660

```

```

RESULT 4
BX385751
LOCUS      717 bp mRNA linear EST 08-MAY-2001
DEFINITION Homo sapiens CDNA clone IMAGE:4822894 5',
            mRNA sequence.
ACCESSION  BX385751
VERSION     BX385751.1 GI:13998422
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            NIH-MGC http://mgs.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshituki and Piero Carninci (RIKEN)
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            CDNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM10731 row: o column: 23
            High quality sequence stop: 708.
            Location/Qualifiers
              1. .717
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"

```



```
QY 549 TTTCATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCGAGCTATTTCAAAGT 608
DB 541 TTTCATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCGAGCTATTTCAAAGT 600
QY 609 GTGTTGGATTAAATTAGGATCATCCCTTGTGTTAAATAAATAAATGCTTTGTGCT 662
DB 601 GTGTTGGATTAAATTAGGATCATCCCTTGTGTTAAATAAATAAATGCTTAGTGCT 654

RESULT 6
AV727829 786 bp mRNA linear EST 17-OCT-2000
AV727829 HTC Homo sapiens cDNA clone HTCCMG07 5', mRNA sequence.
AV727829
AV727829.1 GI:10837250
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L.,
Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
Chen,J., Chen,Z. and Han,Z.
Homo sapiens cDNA HTC clones
Unpublished (2000)
CONTACT Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
LOCATION/Qualifiers
1..786
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HTCCMG07"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_lib="HTC"
/vector="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match 97.7%; Score 647; DB 1; Length 786;
Best Local Similarity 99.2%; Pred. No. 2e-146;
Matches 650; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 8 AGCGAGACCCCTCTCTGCGAGCGCCGCCGCCGCCGCCGCCGCCAGTTCAGCAGC 67
DB 1 AGCGAGACCCCTCTCTGCGAGCGCCGCCGCCGCCGCCGCCGCCAGTTCAGCAGC 60
QY 68 TCGAAGGAGAGTGGCCCTGTGGTGGACGACCAAGGCTTTTCATGATACATGAAGGAGCTAG 127
DB 61 TCGAAGGAGAGTGGCCCTGTGGTGGACGACCAAGGCTTTTCATGATACATGAAGGAGCTAG 120
QY 128 GAGTGGGAATAGCTTTTGGCAAAATGGCGCAATGGCGCAATGGCGCAATGGCGCAATGGCGCAAT 187
DB 121 GAGTGGGAATAGCTTTTGGCAAAATGGCGCAATGGCGCAATGGCGCAATGGCGCAATGGCGCAAT 180
QY 188 GTGATGGTAAAACTTCACATAAACTGAGAGCACTTTGAAACCAACACAGATTTTCTT 247
DB 181 GTGATGGTAAAACTTCACATAAACTGAGAGCACTTTGAAACCAACACAGATTTTCTT 240
QY 248 GTACCTGGGAGAGTTTGAAGAACCAACAGCTGATGCGAGAAACTCAGACTGTCT 307
DB 241 GTACCTGGGAGAGTTTGAAGAACCAACAGCTGATGCGAGAAACTCAGACTGTCT 300
```

```
QY 308 GCAACTTTACAGATGGTGCATTTGGTTTACAGATCAGAGTGGATGGGAAGGAAGCACAA 367
DB 301 GCAACTTTTACAGATGGTGCATTTGGTTTACAGATCAGAGTGGATGGGAAGGAAGCACAA 360
QY 368 TAAACAAGAAAATTGAAGATGGAAAATTAGTGGTGGAGTGTGTGTCATGAACAATGTCACT 427
DB 361 TAAACAAGAAAATTGAAGATGGAAAATTAGTGGTGGAGTGTGTGTCATGAACAATGTCACT 420
QY 428 GTACTCGGATCTATGAAAAGTAGAATAAAAAATTCATCATCATCTTTTGGACAGGAGTTAA 487
DB 421 GTACTCGGATCTATGAAAAGTAGAATAAAAAATTCATCATCATCTTTTGGACAGGAGTTAA 480
QY 488 TTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCATCTTTCTTTT 547
DB 481 TTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCATCTTTCTTTT 540
QY 548 TTTTCAATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCGAGCTATTTCAAAG 607
DB 541 CTTTCAATTACTGTGTTCAATTATCTTTATCATAAACATTTTACATGCGAGCTATTTCAAAG 600
QY 608 TGTGTTGGATTAAATTAGGATCATCCCTTGTGTTAAATAAATAAATGTTTGTCT 662
DB 601 TGTGTTGGATTAAATTAGGATCATCCCTTGTGTTAAATAAATAAATGTTTGTCT 655

RESULT 7
BG715536 708 bp mRNA linear EST 08-MAY-2001
602675842F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4798402 5',
mRNA sequence.
BG715536
BG715536.1 GI:13994723
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 708)
AUTHORS NIH-MGC http://mgc.mci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10686 row: c column: 11
High quality sequence stop: 687.
LOCATION/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798402"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/vector="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

FEATURES
source
LOCATION/Qualifiers
1..708
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4798402"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/vector="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN
```


Db 500 GTTAATTAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTCTTTCT 559
 Qy 543 TTTTCTTTTCAATCTGCTGCTCAATTAATCTTTATCATAAACATTTTACATGCGAGCTATTT 602
 Db 560 TTTTCTTTTCAATCTGCTGCTCAATTAATCTTTATCATAAACATTTTACATGCGAGCTATTT 619
 Qy 603 CAAAGTGTGCTGCTGCTCAATTAATGAGTATCCCTTTGTTTAAATAAATGCTTTGTGC 661
 Db 620 CAAAGTGTGCTGCTGCTCAATTAATGAGTATCCCTTTGTTTAAATAAATGCTTTGTGC 677

RESULT 9
 BM789868 666 bp mRNA linear EST 05-MAR-2002
 LOCUS K-EST0069550 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-8-D02 5',
 DEFINITION mRNA sequence.
 ACCESSION BM789868
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
 Kim, Y.S.
 TITLE 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 8 row: D column: 02
 High quality sequence stop: 666.
 Location/Qualifiers
 1..666
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="S21SNUS20-8-D02"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /lab_host="Top10F"
 /clone_lib="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

FEATURES

source
 1..666
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="S21SNUS20-8-D02"
 /sex="F"
 /tissue_type="Stomach"
 /cell_type="Floating aggregates"
 /lab_host="Top10F"
 /clone_lib="S21SNUS20"
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then decapped
 with tobacco acid pyrophosphatase (TAP). The decapped
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10F by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Query Match 97.1%; Score 643; DB 4; Length 666;
 Best Local Similarity 99.8%; Pred. No. 1.8e-145;
 Matches 654; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

FEATURES

Location/Qualifiers

Qy 3 CCGCAGCGCAGACCCCTCTCTGCAGCGCCAGCCCGCCGACCCACCATGGCCACAGTTCA 62
 Db 13 CCGCAGCGCAGACCCCTCTCTGCAGCGCCAGCCCGCCGACCCACCATGGCCACAGTTCA 72
 Qy 63 CGAGCTGGAAGGAGATGGCGCTGCTGGACAGCAAGGCTTTGATGAATACATGAAGA 122
 Db 73 CGAGCTGGAAGGAGATGGCGCTGCTGGACAGCAAGGCTTTGATGAATACATGAAGA 132
 Qy 123 GCTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCAAAGCCAGATGTATCAT 182
 Db 133 GCTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCAAAGCCAGATGTATCAT 192
 Qy 183 CACTTGTGATGCTTAAACCTCACCATAAATCTGAGAGCACTTTGAAAAACACACAGTT 242
 Db 193 CACTTGTGATGCTTAAACCTCACCATAAATCTGAGAGCACTTTGAAAAACACACAGTT 252
 Qy 243 TTCTTGTACCTCTGGGAGAGAGTGTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 302
 Db 253 TTCTTGTACCTCTGGGAGAGAGTGTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 312
 Qy 303 TGTCTGCAACTTTACAGATGGTGCAATGGTTCAGATCAGGAGTGGGATGGGAAGAAAG 362
 Db 313 TGTCTGCAACTTTACAGATGGTGCAATGGTTCAGATCAGGAGTGGGATGGGAAGAAAG 372
 Qy 363 CACATAACAGAAAAATTTGAAGATGGGAAATTTAGTGGTGGTGTCTCATGAACAATGT 422
 Db 373 CACATAACAGAAAAATTTGAAGATGGGAAATTTAGTGGTGGTGTCTCATGAACAATGT 432
 Qy 423 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCTTTGGACAGGA 482
 Db 433 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCTTTGGACAGGA 492
 Qy 483 GTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCATCTTTCTTCT 542
 Db 493 GTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCATCTTTCTTCT 552
 Qy 543 TTTTCTTTTTCATTCTGTTTCAATATCTTTTATATAAAGATTTTACATGCGAGCTATTT 602
 Db 553 TTTTCTTTTTCATTCTGTTTCAATATCTTTTATATAAAGATTTTACATGCGAGCTATTT 612
 Qy 603 CAAAGTGTGCTGATTAATAGGATCATCCCTTTGTTTAAATAAATGCTTTT 657
 Db 613 CAAAGTGTGCTGATTAATAGGATCATCCCTTTGTTTAAATAAATGCTTTT 666

RESULT 10

LOCUS BU601920
 DEFINITION BU601920 843 bp mRNA linear EST 20-SEP-2002
 IMAGE:6495933 5', mRNA sequence.
 ACCESSION BU601920
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 843)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue procurement: NCI
 cDNA library preparation: Michael Brownstein Laboratory
 cDNA library arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM2674 row: 1 column: 22
 High quality sequence stop: 542.
 Location/Qualifiers

source

```
1. .843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6495933"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phrase-resistant)"
/clone_lib="NIH_MGC_142"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccgcctccggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGAGTGGCCATTACGGCGGG-3' and
5'-ATTCTAGAGCGGCGGCCGACATG-DT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the >0.5 kb
size fraction (other fractions present in NIH_MGC_141).
Library created in the laboratory of M. Brownstein (NIMH,
NIH). Note: this is a NIH_MGC Library."
```

ORIGIN

```
Query Match          96.8%; Score 640.8; DB 5; Length 843;
Best Local Similarity 98.2%; Pred. No. 6.5e-145;
Matches 648; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 CGCCGACGAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATCGGCCACAGTTCA 62
DB 18 CGCCGACGAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATCGGCCACAGTTCA 77

QY 63 GCAGCTGGAAGGAAGATGCGCCTCTGTGACAGCAAGAGCTTTGATGATACATGAAGGA 122
DB 78 GCAGCTGGAAGGAAGATGCGCCTCTGTGACAGCAAGAGCTTTGATGATACATGAAGGA 137

QY 123 GCTAGAGTGGGAATAGCTTTTCGGAATATGGCCCAATGGCCCAATGGCCCAATGGTATCAT 182
DB 138 GCTAGAGTGGGAATAGCTTTTCGGAATATGGCCCAATGGCCCAATGGTATCAT 197

QY 183 CACTTGTGATGTAAACCACTCACCATAAACTGAGAGCACTTTGAAAAACCAACAGTT 242
DB 198 CACTTGTGATGTAAACCACTCACCATAAACTGAGAGCACTTTGAAAAACCAACAGTT 257

QY 243 TTCCTGTACCTCGGAGAGAGTTTGAAGAAACCAACAGCTGATGCGAGAAACCTCAGAC 302
DB 258 TTCCTGTACCTCGGAGAGAGTTTGAAGAAACCAACAGCTGATGCGAGAAACCTCAGAC 317

QY 303 TGCTGCACTTTTACAGATGGTGCATTTGGTTTCAGCATCAGGATGGGATGGGAAGGAAG 362
DB 318 TGCTGCACTTTTACAGATGGTGCATTTGGTTTCAGCATCAGGATGGGATGGGAAGGAAG 377

QY 363 CACAATAACAGAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTGATGAACAATGT 422
DB 378 CACAATAACAGAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTGATGAACAATGT 437

QY 423 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCATCTTTGGACAGGA 482
DB 438 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCATCTTTGGACAGGA 497

QY 483 GTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAATCTCCATCTGTTCTTTCT 542
DB 498 GTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAATCTCCATCTGTTCTTTCT 557

QY 543 TTTTCTTTTTCATTACTGTGTTCAATTAATCTTTTATCATAAACATTTTATATGACAGCTATTT 602
DB 558 TTTTCTTTTTCATTACTGTGTTCAATTAATCTTTTATCATAAACATTTTATATGACAGCTATTT 617

QY 603 CAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTGGTAAATAAATAATGTTGTGCT 662
DB 618 CAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTGGTAAATAAATAATGTTGTGCT 677
```

RESULT 11

```
AV724494
LOCUS          AV724494
DEFINITION    HTB Homo sapiens CDNA clone HTBAYF07 5', mRNA sequence.
ACCESSION     AV724494
VERSION       AV724494.1 GI:10828936
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 656)
AUTHORS      Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
              Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L.,
              Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
              Chen,J., Chen,Z. and Han,Z.
              Homo sapiens CDNA HTB clones
              Unpublished (2000)
              Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919 (ex.45)
              Fax: 86-21-50801922
              Email: hanzg@chgc.sh.cn
              This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source           1..656
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HTBAYF07"
                     /tissue_type="Hypothalamus"
                     /dev_stage="Adult"
                     /lab_host="SOLR"
                     /clone_lib="HTB"
                     /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                     XhoI"
ORIGIN
Query Match          96.7%; Score 640.4; DB 1; Length 656;
Best Local Similarity 99.8%; Pred. No. 7.8e-145;
Matches 641; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GCGGACGAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATCGGCCACAGTTCA 63
DB 1 GCGGACGAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATCGGCCACAGTTCA 60

QY 64 CAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAGGCTTTGATGAATACATGAAGGAG 123
DB 61 CAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAGGCTTTGATGAATACATGAAGGAG 120

QY 124 CTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCCAAGCCAGATTGTATCATC 183
DB 121 CTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCCAAGCCAGATTGTATCATC 180

QY 184 ACTTGTGATGTAAACCTCACCATAAACTGAGAGCACTTTGAAACCAACACAGTTT 243
DB 181 ACTTGTGATGTAAACCTCACCATAAACTGAGAGCACTTTGAAACCAACACAGTTT 240

QY 244 TCTTGTACCTCGGAGAGAGCTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGACT 303
DB 241 TCTTGTACCTCGGAGAGAGCTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGACT 300

QY 304 GTCTGCAACTTTTACAGATGGTGCAATGGTTTTCAGCATCAGGATGGGATGGGAAGAAAGC 363
DB 301 GTCTGCAACTTTTACAGATGGTGCAATGGTTTTCAGCATCAGGATGGGATGGGAAGAAAGC 360

QY 364 ACAATAACAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTCTCATGAACATGTC 423
DB 361 ACAATAACAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTCTCATGAACATGTC 420
```



```
QY 303 TGTCTGCAACTTTACAGATGGTGCATTGGTTTCAGCATCAGGAGTGGGATGGGAAGGAAG 362
Db |||||
QY 317 TGTCTGCAACTTTACAGATGGTGCATTGGTTTCAGCATCAGGAGTGGGATGGGAAGGAAG 376
Db |||||
QY 363 CACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTCATCAACAATGT 422
Db |||||
QY 377 CACATTAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTCATCAACAATGT 436
Db |||||
QY 423 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCACTTTTGGACAGGA 482
Db |||||
QY 437 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCACTTTTGGACAGGA 496
Db |||||
QY 483 GTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCCCATACTGTTCTTCT 542
Db |||||
QY 497 GTTAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCCCATACTGTTCTTCT 556
Db |||||
QY 543 TTTTCTTTTTCATTACTGTGTTCAATTAATCTTTTATCATAAACATTTTACATGCAGCTATTT 602
Db |||||
QY 557 TTTTCTTTTTCATTACTGTGTTCAATTAATCTTTTATCATAAACATTTTACATGCAGCTATTT 616
Db |||||
QY 603 CAAAGTGT-GTTGGATTAATTAGGATCAT-CCCTTTGGTTAATAAATAAATGTGTTTG 658
Db |||||
QY 617 CAAAGTGTGTTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTG 674
```

Search completed: July 12, 2005, 21:11:53
Job time : 2590.24 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	662	100.0	662	6	CQ776686	Sequence
2	662	100.0	662	6	CQ861699	Sequence
3	662	100.0	662	6	AX224147	Sequence
4	662	100.0	662	9	HUMFABPHA	Human fatty
5	660	99.7	661	6	CQ719340	Sequence
6	660	99.7	680	9	BC019385	Homo sapi
7	660	99.7	720	6	BD203735	Human nuc
8	660	99.7	720	6	AC014890	Sequence
9	658	99.4	706	9	BC070303	Homo sapi
10	649	98.0	1072	6	AX281734	Sequence
11	645	97.4	645	6	CQ729706	Sequence
C 12	645	97.4	156734	9	AP000640	Homo sapi
C 13	645	97.4	180726	2	AC090343	Homo sapi
C 14	645	97.4	195396	2	AC021809	Homo sapi
15	642.4	97.0	661	6	CQ725017	Sequence
C 16	625.6	94.5	199882	9	AL354720	Human DNA
17	623.2	94.1	194150	2	AC148825	Pan trogl
18	618.6	93.4	651	6	CQ724062	Sequence
C 19	618.4	93.4	194575	2	AC023140	Homo sapi

QY 121 GAGCTAGGAGTGGGAATAGCTTTTCGGAATAATGGGGCGCAATGGCCAGCCAGATTGTATC 180
 |||||
 Db 121 GAGCTAGGAGTGGGAATAGCTTTTCGGAATAATGGGGCGCAATGGCCAGCCAGATTGTATC 180
 |||||
 QY 181 ATCACTTGTGATGGTAAACCTTACCACATATAAAGTGGAGGACCTTTGAAAAACACACAG 240
 |||||
 Db 181 ATCACTTGTGATGGTAAACCTTACCACATATAAAGTGGAGGACCTTTGAAAAACACACAG 240
 |||||
 QY 241 TTTTCTTGTAACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGCGAGAAACCTCAG 300
 |||||
 Db 241 TTTTCTTGTAACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGCGAGAAACCTCAG 300
 |||||
 QY 301 ACTGTCGCACTTTACAGATGTCGATTTGGTTACGATCAGGAGTGGAGTGGGAAGGAA 360
 |||||
 Db 301 ACTGTCGCACTTTACAGATGTCGATTTGGTTACGATCAGGAGTGGAGTGGGAAGGAA 360
 |||||
 QY 361 AGCACAATAACAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTGATGAACAAT 420
 |||||
 Db 361 AGCACAATAACAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTGATGAACAAT 420
 |||||
 QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
 |||||
 Db 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
 |||||
 QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTT 540
 |||||
 Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTT 540
 |||||
 QY 541 CTTTCTTTTTCATTAAGTGTGTTCAATTAATCTTTTATCAATAAATTTTACATGACGCTAT 600
 |||||
 Db 541 CTTTCTTTTTCATTAAGTGTGTTCAATTAATCTTTTATCAATAAATTTTACATGACGCTAT 600
 |||||
 QY 601 TTCAAGTGTGTTGGATTAATAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 |||||
 Db 601 TTCAAGTGTGTTGGATTAATAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 |||||
 QY 661 CT 662
 ||
 Db 661 CT 662

RESULT 4
 HUMFABPHA
 LOCUS 662 bp mRNA linear PRI 31-DEC-1994
 DEFINITION Human fatty acid binding protein homologue (PA-FABP) mRNA, complete cds.

ACCESSION M94856
 VERSION M94856.1 GI:182353
 KEYWORDS fatty acid binding protein homologue.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Madesen, P., Rasmussen, H.H., Leffers, H., Honore, B. and Celis, J.E.
 TITLE Molecular cloning and expression of a novel keratinocyte protein (psoriasis-associated fatty acid-binding protein [PA-FABP]) that is highly up-regulated in psoriatic skin and that shares similarity to fatty acid-binding proteins
 J. Invest. Dermatol. 99 (3), 299-305 (1992)

JOURNAL MEDLINE 92381332

COMMENT Original source text: Homo sapiens (tissue library: lambda gt11)
 adult epidermis cDNA to mRNA.

FEATURES
 source Location/Qualifiers

1..662
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="17"
 /cell_line="unfractionated non-cultured keratinocyte"
 /cell_type="keratinocyte"

RESULT 5
 CQ719340

gene /tissue_type="epidermis"
 /dev_stage="adult"
 /tissue_lib="lambda gt11"
 1..662
 /gene="PA-FABP"
 49..456
 /gene="PA-FABP"
 /codon_start=1
 /product="fatty acid binding protein homologue"
 /protein_id="AA58467.1"
 /db_xref="GI:182353"
 /translation="MATVQQLGRWFLVDSKGFDEYKELGVGIALRKMCAKAPDCI
 ITCGKXLIKTESTLKTTFQFSCITLGEKEFEETADGRKTQVCNFDGALVQHQWDG
 KESTITRKUDGKLWVECVNNNVCTRIYKVE"
 polyA_signal 645..650
 /gene="PA-FABP"
 ORIGIN chromosome 17.

Query Match 100.0%; Score 662; DB 9; Length 662;
 Best Local Similarity 100.0%; Pred. No. 1.7e-146;
 Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCGCCGACGACAGCCCTCTCTGCACGCGCAGCCCGCCGCCACCCATGGCCACAGTT 60
 Db 1 ACCGCCGACGACAGCCCTCTCTGCACGCGCAGCCCGCCGCCACCCATGGCCACAGTT 60
 QY 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGAGCAAGCAAGGCTTTGATGAATACATGAAG 120
 Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGAGCAAGCAAGGCTTTGATGAATACATGAAG 120
 QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGGGCGCAATGGCCAGCCAGATTGTATC 180
 Db 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGGGCGCAATGGCCAGCCAGATTGTATC 180
 QY 181 ATCACTTGTGATGGTAAACCTTACCACATATAAAGTGGAGGACCTTTGAAAAACACACAG 240
 Db 181 ATCACTTGTGATGGTAAACCTTACCACATATAAAGTGGAGGACCTTTGAAAAACACACAG 240
 QY 241 TTTTCTTGTAACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGCGAGAAACCTCAG 300
 Db 241 TTTTCTTGTAACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGCGAGAAACCTCAG 300
 QY 301 ACTGTCGCACTTTACAGATGTCGATTTGGTTACGATCAGGAGTGGAGTGGGAAGGAA 360
 Db 301 ACTGTCGCACTTTACAGATGTCGATTTGGTTACGATCAGGAGTGGAGTGGGAAGGAA 360
 QY 361 AGCACAATAACAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTGATGAACAAT 420
 Db 361 AGCACAATAACAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTGATGAACAAT 420
 QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
 Db 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
 QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTT 540
 Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTT 540
 QY 541 CTTTCTTTTTCATTAAGTGTGTTCAATTAATCTTTTATCAATAAATTTTACATGACGCTAT 600
 Db 541 CTTTCTTTTTCATTAAGTGTGTTCAATTAATCTTTTATCAATAAATTTTACATGACGCTAT 600
 QY 601 TTCAAGTGTGTTGGATTAATAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 Db 601 TTCAAGTGTGTTGGATTAATAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 QY 661 CT 662
 ||
 Db 661 CT 662

LOCUS CQ719340 661 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence S274 from Patent WO02068579.
ACCESSION CQ719340
VERSION CQ719340.1 GI:42280197
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kits, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 5274 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. 661
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.7%; Score 660; DB 6; Length 661;
Best Local Similarity 100.0%; Pred. No. 5.2e-146;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CGCCGACGAGACCCCTCTCTGCAGCCAGCCGCCGCCGCCACCCACCATGCGCCACAGTTCA 62
Db 2 CGCCGAGCGAGACCCCTCTCTGCAGCCAGCCGCCGCCGCCACCCACCATGCGCCACAGTTCA 61
QY 63 GCAGCTGGAAGGAAGATGCGCGCTGTGTGACGACGAAAGCTTTGATGATATCATGAGGA 122
Db 62 GCAGCTGGAAGGAAGATGCGCGCTGTGTGACGACGAAAGCTTTGATGATATCATGAGGA 121
QY 123 GCTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 182
Db 122 GCTAGAGTGGGAATAGCTTTTCGAAAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 181
QY 183 CACTTGTGTGTTAAACCCCTCACCATAAAACCTGAGAGCACTTTGAAACCAACACAGTT 242
Db 182 CACTTGTGTGTTAAACCCCTCACCATAAAACCTGAGAGCACTTTGAAACCAACACAGTT 241
QY 243 TTCTTGTACCTGGGAGAGAAGTTTGAAGAAACCCAGCTGATGGCAGAAAACCTCAGAC 302
Db 242 TTCTTGTACCTGGGAGAGAAGTTTGAAGAAACCCAGCTGATGGCAGAAAACCTCAGAC 301
QY 303 TGTCTGCACTTTACAGATGGTGCATTGGTTGAGCAATCAGGAGTGGGATGGGAGGAAG 362
Db 302 TGTCTGCACTTTACAGATGGTGCATTGGTTGAGCAATCAGGAGTGGGATGGGAGGAAG 361
QY 363 CACAATAACAAGAAATTAAGAGATGGGAAATTTAGTGGTGGAGTGTGTCATGAACAATGT 422
Db 362 CACAATAACAAGAAATTAAGAGATGGGAAATTTAGTGGTGGAGTGTGTCATGAACAATGT 421
QY 423 CACCTGTACTCGATCTATGAAAGATAGAAATAAAATTCATCATCTCTTTGGACAGGA 482
Db 422 CACCTGTACTCGATCTATGAAAGATAGAAATAAAATTCATCATCTCTTTGGACAGGA 481
QY 483 GTTAATTAAGAGAAATCAGCAAGCTCAGTTCAATGAGCAAAATCTCCATATCTGTTCTTCT 542
Db 482 GTTAATTAAGAGAAATCAGCAAGCTCAGTTCAATGAGCAAAATCTCCATATCTGTTCTTCT 541
QY 543 TTTTCTTTTTCATTACTGTGTCAATTTATCTTTATCATAAACATTTTACATGACGATTTT 602
Db 542 TTTTCTTTTTCATTACTGTGTCAATTTATCTTTATCATAAACATTTTACATGACGATTTT 601
QY 603 CAAAGTGTGTGGATTAATTAGCATCATCCCTTGTGTTAAATAAATGCTTTGTGCT 662
Db 602 CAAAGTGTGTGGATTAATTAGCATCATCCCTTGTGTTAAATAAATGCTTTGTGCT 661

RESULT 6

BC019385
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source
gene

BC019385 680 bp mRNA linear PRI 29-JUN-2004
Homo sapiens fatty acid binding protein 5 (psoriasis-associated),
mRNA (cDNA clone MGC:9003 IMAGE:3862100), complete cds.
BC019385
BC019385.2 GI:34783324
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 680)
Straube, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
Sanchez, A., Whitting, M., Madan, A., Touchman, J.W., Green, E.D.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallov, D.E.,
Schnur, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 680)
Straube, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Sep 16, 2003 this sequence version replaced gi:18043065.
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunnar, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louie, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 21 Row: 9 Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4557580.
Location/Qualifiers
1. .680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:9003 IMAGE:3862100"
/tissue_type="Ovary, adenocarcinoma"
/clone_lib="NIH_MGC_66"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .680

/gene="FABPS"
/note="synonyms: E-FABP, PAFABP, EFABP, PA-FABP"
/db_xref="LocusID:2171"
/db_xref="MIM:605168"
47_454
/gene="FABPS"
/codon_start=1
/product="fatty acid binding protein 5
(seriolysis-associated)"
/protein_id="AAH19385.1"
/db_xref="GI:18043066"
/db_xref="LocusID:2171"
/db_xref="MIM:605168"
translation="MATVOOLEGRWRLVDSKGFDEYKELGVIALRKNGAMAKPDCI
ITCDGNLRIKSTLTKTQFSCITLGEKEFTTADGRKTQTVNFTDGLVQHQEWDC
KESTIRKLKDGKLVVECMNVNVTCTRIYKVE"

ORIGIN

Query Match 99.7%; Score 660; DB 9; Length 680;
Best Local Similarity 100.0%; Pred. No. 5.2e-146;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCACCATGGCCACAGTTCA 62
DB 1 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCACCATGGCCACAGTTCA 60
QY 63 GCAGCTGGAAGGAGATGGCGCTGTGGACAGCAAGCGCTTTGATGAATACATGAAGGA 122
DB 61 GCAGCTGGAAGGAGATGGCGCTGTGGACAGCAAGCGCTTTGATGAATACATGAAGGA 120
QY 123 GCTAGAGTGGGAATAGCTTTGCGAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 182
DB 121 GCTAGAGTGGGAATAGCTTTGCGAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 180
QY 183 CACTTGTGATGTGTAAGAACTCACCATAAAGCTGAGACACTTTGAAACACACAGTT 242
DB 181 CACTTGTGATGTGTAAGAACTCACCATAAAGCTGAGACACTTTGAAACACACAGTT 240
QY 243 TTCTTTTACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 302
DB 241 TTCTTTTACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 300
QY 303 TGTCTGCAACTTTACAGATGGTGCATTTGATCAGCATCAGGAGTGGGAGGGAAG 362
DB 301 TGTCTGCAACTTTACAGATGGTGCATTTGATCAGCATCAGGAGTGGGAGGGAAG 360
QY 363 CACAAATAACAGAAATTCAGATGGGAATTTAGTGTGGAGTGTCTATGAACAATCT 422
DB 361 CACAAATAACAGAAATTCAGATGGGAATTTAGTGTGGAGTGTCTATGAACAATCT 420
QY 423 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 482
DB 421 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 480
QY 483 GTTAATTAAGAGAAATCAAGCTCAGTTCAATGAGCAAAATCTCCATCTCTTTCTTCT 542
DB 481 GTTAATTAAGAGAAATCAAGCTCAGTTCAATGAGCAAAATCTCCATCTCTTTCTTCT 540
QY 543 TTTTCTTTTTCATTTACTGTGTTCAATTTATCTTTATCATAAACATTTTACATGACGCTATTT 502
DB 541 TTTTCTTTTTCATTTACTGTGTTCAATTTATCTTTATCATAAACATTTTACATGACGCTATTT 500
QY 603 CAAAGTGTGTGGATTAATTTAGGATCATCTCTTTGGTTAATAAATAAATGTGTTGTGCT 562
DB 601 CAAAGTGTGTGGATTAATTTAGGATCATCTCTTTGGTTAATAAATAAATGTGTTGTGCT 560

RESULT 7

BD203735 720 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Human nucleic acid sequence originating in ovarian myoma tissue.
ACCESSION BD203735
VERSION BD203735.1 GI:33013505

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JP 2002511252-A/76.
Homo sapiens (human)
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 720)
Specht, F., Hinzmann, B., Schmitt, A., Pilarsky, C., Dahl, E. and Rosenthal, A.
Human nucleic acid sequence originating in ovarian myoma tissue
Patent: JP 2002511252-A 76 16-APR-2002;
METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
OS Homo sapiens (human)
PN JP 2002511252-A/76
PD 16-APR-2002
PR 07-APR-1999 JP 2000543588
PR 09-APR-1998 DE 198 17 557.4
PI THOMAS SPECHT, BERND HINZMANN, ARMIN SCHMITT, CHRISTIAN PILARSKY,
PI EDGAR DAHL,
PI ANDRE ROSENTHAL
PC C12N15/09, A61K38/00, A61P15/00, A61P35/00, C07K14/82,
PC C07K16/32,
PC C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/08, C12Q1/68, G01N33/574,
PC C12N15/00, A61K37/02, C12N5/00
CC Human nucleic acid sequence originating in ovarian myoma CC

FEATURES

source

Location/Qualifiers

1..720

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 99.7%; Score 660; DB 6; Length 720;

Best Local Similarity 100.0%; Pred. No. 5.2e-146;

Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCACCATGGCCACAGTTCA 62
DB 42 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCACCATGGCCACAGTTCA 101
QY 63 GCAGCTGGAAGGAGATGGCGCTGTGGACAGCAAGCGCTTTGATGAATACATGAAGGA 122
DB 102 GCAGCTGGAAGGAGATGGCGCTGTGGACAGCAAGCGCTTTGATGAATACATGAAGGA 161
QY 123 GCTAGAGTGGGAATAGCTTTGCGAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 182
DB 162 GCTAGAGTGGGAATAGCTTTGCGAAATGGCGCAATGGCCAGCCAGATTTGTATCAT 221
QY 183 CACTTGTGATGTGTAAGAACTCACCATAAAGCTGAGACACTTTGAAACACACAGTT 242
DB 222 CACTTGTGATGTGTAAGAACTCACCATAAAGCTGAGACACTTTGAAACACACAGTT 281
QY 243 TTCTTTTACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 302
DB 282 TTCTTTTACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCAGAAAACTCAGAC 341
QY 303 TGTCTGCAACTTTACAGATGGTGCATTTGATCAGCATCAGGAGTGGGAGGGAAG 362
DB 342 TGTCTGCAACTTTACAGATGGTGCATTTGATCAGCATCAGGAGTGGGAGGGAAG 401
QY 363 CACAAATAACAGAAATTCAGATGGGAATTTAGTGTGGAGTGTCTATGAACAATCT 422
DB 402 CACAAATAACAGAAATTCAGATGGGAATTTAGTGTGGAGTGTCTATGAACAATCT 461
QY 423 CACTTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 482
DB 462 CACTTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGA 521
QY 483 GTTAATTAAGAGAAATCAAGCTCAGTTCAATGAGCAAAATCTCCATCTCTTTCTTCT 542

```
|||||
Db 522 GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTCTTCT 581
|||||
Qy 543 TTTTCTTTTTCATTACTGTGTTCAATATCTTTTATCATAAACATTTTACATGAGCTATTT 602
|||||
Db 582 TTTTCTTTTTCATTACTGTGTTCAATATCTTTTATCATAAACATTTTACATGAGCTATTT 641
|||||
Qy 603 CAAAGTGTGTTGGATTAATTAGATCATCCCTTTGGTTTAATAAATAAATGTTTGTGCT 662
|||||
Db 642 CAAAGTGTGTTGGATTAATTAGATCATCCCTTTGGTTTAATAAATAAATGTTTGTGCT 701

RESULT 8
AX014890
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
Human nucleic acid sequences from ovarian tumour tissue
Patent: WO 9953040-A 89 21-OCT-1999;
SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GBS FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
Location/Qualifiers
1. .720
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

FEATURES
source
ORIGIN
Query Match 99.7%; Score 660; DB 6; Length 720;
Best Local Similarity 100.0%; Pred. NO. 5.2e-146;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCATGCGCACAGTTCA 62
Db 42 CGCCGACGACAGCCCTCTCTGACGCCAGCCCGCCGACCCACCATGCGCACAGTTCA 101
Qy 63 CGAGCTGGAAGGAAGTGGCGCTGTGTGGACAGCAAGGCTTTGATGATATACATGAAGA 122
Db 102 CGAGCTGGAAGGAAGTGGCGCTGTGTGGACAGCAAGGCTTTGATGATATACATGAAGA 161
Qy 123 GCTAGAGTGGGAATAGCTTTTGGGAAAATGGCGCAATGGCCAGCCAGATTGTATCAT 182
Db 162 GCTAGAGTGGGAATAGCTTTTGGGAAAATGGCGCAATGGCCAGCCAGATTGTATCAT 221
Qy 183 CACTTGTGATGTGTAAACCTCACCATAAAATCTGAGAGCACTTTGAAAACCAACAGTT 242
Db 222 CACTTGTGATGTGTAAACCTCACCATAAAATCTGAGAGCACTTTGAAAACCAACAGTT 281
Qy 243 TTCTTTGACCTGGGAGAGAAGTTTGAAGAAACCAAGCTGATGGCAGAAAACCTCAGAC 302
Db 282 TTCTTTGACCTGGGAGAGAAGTTTGAAGAAACCAAGCTGATGGCAGAAAACCTCAGAC 341
Qy 303 TGTCTGCACTTTACAGATGGTGCAATGGTTGATCATAGGAGTGGGATGGGAAGGAAG 362
Db 342 TGTCTGCACTTTACAGATGGTGCAATGGTTGATCATAGGAGTGGGATGGGAAGGAAG 401
Qy 363 CACAATAACAAGAAATTCGAAGATGGGAAATTTAGTGTGGAGTGTGTCATGAACAATCT 422
Db 402 CACAATAACAAGAAATTCGAAGATGGGAAATTTAGTGTGGAGTGTGTCATGAACAATCT 461
Qy 423 CACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCTTTGGACAGGA 482
Db 462 CACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCTTTGGACAGGA 521
```

```
Qy 483 GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTCTTCT 542
Db 522 GTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTCTTCT 581
Qy 543 TTTTCTTTTTCATTACTGTGTTCAATATCTTTTATCATAAACATTTTACATGAGCTATTT 602
Db 582 TTTTCTTTTTCATTACTGTGTTCAATATCTTTTATCATAAACATTTTACATGAGCTATTT 641
Qy 603 CAAAGTGTGTTGGATTAATTAGATCATCCCTTTGGTTTAATAAATAAATGTTTGTGCT 662
Db 642 CAAAGTGTGTTGGATTAATTAGATCATCCCTTTGGTTTAATAAATAAATGTTTGTGCT 701
```

```
RESULT 9
BC070303
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 706)
Straussberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Loquellano,N.A., Peters,G.J.,
Carninci,P., Prange,C., Raha,S.S., Bosak,S.A., McEwan,P.J.,
Abramson,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
McKernan,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shewchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
```

```
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 706)
Straussberg,R.
```

```
Direct Submission
Submitted (10-MAY-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
```

```
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
```

```
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
```

```
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAL Plate: 57 Row: j Column: 10
```



```
|||||
582 TTTTTCATTTTCACTGTTCAATATCTTTATCATAAACATTTTACATGCAGCTATT 641
|||||
602 TCAAGTGTGTGGATTAATTAAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGGC 661
|||||
642 TCAAGTGTGTGGATTAATTAAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGGC 701
|||||
662 T 662
702 T 702

RESULT 11
CQ729706
LOCUS CQ729706 645 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15640 from Patent WO02069579.
ACCESSION CQ729706
VERSION CQ729706.1 GI:42301523
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humaneons or transcripts, for detecting expression and other uses
JOURNAL Patent: WO 02068579-A 15640 06-SEP-2002;
PE Corporation (NV) (US)
FEATURES
source
Location/Qualifiers
1..645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 97.4%; Score 645; DB 6; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.9e-142; Indels 0; Gaps 0;
Matches 645; Conservative 0; Mismatches 0;
1 CTCTCTGCACGCCAGCCCGCCGCCACCCACCATGCGCCACAGTTCACGAGCTGGAAGGAG 77
DB 1 CTCTCTGCACGCCAGCCCGCCGCCACCCACCATGCGCCACAGTTCACGAGCTGGAAGGAG 60
QY 78 ATGGCGCCTGTGTGGACAGCAAAAGGCTTTGATGAATACATGAAGGAGCTAGGAGTGGGAAT 137
DB 61 ATGGCGCCTGTGTGGACAGCAAAAGGCTTTGATGAATACATGAAGGAGCTAGGAGTGGGAAT 120
QY 138 AGCTTTGCGAAAATGGCGCAATGCGCAAGCCAGATTGTATCATCATCTGTGTGATGTTAA 197
DB 121 AGCTTTGCGAAAATGGCGCAATGCGCAAGCCAGATTGTATCATCATCTGTGTGATGTTAA 180
QY 198 AAACCTCACCATAAACCTGAGAGCATTGTGAAAACCAACACAGTTTTCTTTGACCTGGG 257
DB 181 AAACCTCACCATAAACCTGAGGACATTGTGAAAACCAACACAGTTTTCTTTGACCTGGG 240
QY 259 AGAGAAGTTTGAAGAAACACACAGCTGATGCGAGAAAACCTCAGACTGTCTGCAACTTTAC 317
DB 241 AGAGAAGTTTGAAGAAACACACAGCTGATGCGAGAAAACCTCAGACTGTCTGCAACTTTAC 300
QY 318 AGATGTCATTTGGTTTCAGCATCAGGAGTGGGATGGGAGGAAAGCACAATAACAGAAA 377
DB 301 AGATGTCATTTGGTTTCAGCATCAGGAGTGGGATGGGAGGAAAGCACAATAACAGAAA 360
QY 378 ATTGAAAGATGGGAATTAAGTGTGGAGTGTGTGATGAACAATGTCCCTGTACTCGGAT 437
DB 361 ATTGAAAGATGGGAATTAAGTGTGGAGTGTGTGATGAACAATGTCCCTGTACTCGGAT 420
QY 438 CTATGAAAAGTAGAATAAATAATTCATCATCTTTGGACAGGAGTTAATTAAGAGAT 497
DB 421 CTATGAAAAGTAGAATAAATAATTCATCATCTTTGGACAGGAGTTAATTAAGAGAT 480
```

```
QY 498 GACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTCTTTCTTTCTTTTTCATTAC 557
DB |||||||
481 GACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTCTTTCTTTCTTTTTCATTAC 540
|||||
558 TGTGTTCAATTTATCTTTATCATAAACATTTTACATGACGACTATTTCAAAGTGTGTGGAT 617
DB |||||||
541 TGTGTTCAATTTATCTTTATCATAAACATTTTACATGACGACTATTTCAAAGTGTGTGGAT 600
|||||
618 TAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTGCT 662
DB |||||||
601 TAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTGCT 645

RESULT 12
AP000640/c
LOCUS AP000640 156734 bp DNA linear PRI 01-FEB-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11 clone: CMB9-26D16, complete
sequence.
ACCESSION AP000640
VERSION AP000640.5 GI:28189503
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (1999)
REFERENCE 2 (bases 1 to 156734)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (28-OCT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT On Jan 31, 2003 this sequence version replaced gi:17425221.
FEATURES
Location/Qualifiers
1..156734
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="CMB9-26D16"
ORIGIN
Query Match 97.4%; Score 645; DB 9; Length 156734;
Best Local Similarity 100.0%; Pred. No. 1.7e-142; Indels 0; Gaps 0;
Matches 645; Conservative 0; Mismatches 0;
18 CTCTCTGCACGCCAGCCCGCCGCCACCCACCATGCGCCACAGTTCACGAGCTGGAAGGAG 77
DB 121059 CTCTCTGCACGCCAGCCCGCCGCCACCCACCATGCGCCACAGTTCACGAGCTGGAAGGAG 121000
QY 78 ATGGCGCCTGTGTGGACAGCAAAAGGCTTTGATGAATACATGAAGGAGCTAGGAGTGGGAAT 137
DB 120999 ATGGCGCCTGTGTGGACAGCAAAAGGCTTTGATGAATACATGAAGGAGCTAGGAGTGGGAAT 120940
QY 138 AGCTTTGCGAAAATGGCGCAATGCGCAAGCCAGATTGTATCATCATCTGTGTGATGTTAA 197
DB 120939 AGCTTTGCGAAAATGGCGCAATGCGCAAGCCAGATTGTATCATCATCTGTGTGATGTTAA 120880
QY 198 AAACCTCACCATAAACCTGAGAGCATTGTGAAAACCAACACAGTTTTCTTTGACCTGGG 257
DB 120879 AAACCTCACCATAAACCTGAGAGCATTGTGAAAACCAACACAGTTTTCTTTGACCTGGG 120820
QY 258 AGAGAAGTTTGAAGAAACACACAGCTGATGCGAGAAAACCTCAGACTGTCTGCAACTTTAC 317
DB 120819 AGAGAAGTTTGAAGAAACACACAGCTGATGCGAGAAAACCTCAGACTGTCTGCAACTTTAC 120760
```



```
QY 318 AGATGTGTCATTGGTTTCAGCATCAGGAGTGGGATGGGAAGGAAAGCACAATACAGAAA 377
|||||
Db 120759 AGATGTGTCATTGGTTTCAGCATCAGGAGTGGGATGGGAAGGAAAGCACAATACAGAAA 120700
|||||
QY 378 ATTGAAGATGGGAATATTAGTGTGAGTGTGTCATGAACAATGTCACTGTACTCGGAT 437
|||||
Db 120699 ATTGAAGATGGGAATATTAGTGTGAGTGTGTCATGAACAATGTCACTGTACTCGGAT 120640
|||||
QY 438 CTATGAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGAGTAAATTAAGAGAAT 497
|||||
Db 120639 CTATGAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGAGTAAATTAAGAGAAT 120580
|||||
QY 498 GACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTCTTTCTTTTTCATTAC 557
|||||
Db 120579 GACCAAGCTCAGTTCAATGAGCAAAATCTCATACTGTTCTTCTTTCTTTTTCATTAC 120520
|||||
QY 558 TGTGTTCAATTATCTTTATCATAAACATTTTACATGACGATTTTCAAGAGTGTGGAT 617
|||||
Db 120519 TGTGTTCAATTATCTTTATCATAAACATTTTACATGACGATTTTCAAGAGTGTGGAT 120460
|||||
QY 618 TAATTAGGATCATCCCTTTGGTTTAATAAATAAATGTTGTGCT 662
|||||
Db 120459 TAATTAGGATCATCCCTTTGGTTTAATAAATAAATGTTGTGCT 120415
|||||
```

```
RESULT 13
AC090343/c
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-784A24 map 11, WORKING DRAFT
SEQUENCES, 7 unordered pieces.
AC090343
AC090343.2 GI:14971269
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 180726)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-784A24
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 180726)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,I., Bouckgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heathord,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 20, 2001 this sequence version replaced gi:12957979.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
```

```
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12621
Center clone name: 784_A_24
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 176715 bases at least Q40
Consensus quality: 178778 bases at least Q30
Consensus quality: 179540 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 180126; sum-of-contrigs
Quality coverage: 7.4 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 bases; sum-of-contrigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 4557: contig of 4557 bp in length
* 4558 4657: gap of 100 bp
* 4658 5364: contig of 707 bp in length
* 5365 5464: gap of 100 bp
* 5465 7283: contig of 1819 bp in length
* 7284 7383: gap of 100 bp
* 7384 11367: contig of 3984 bp in length
* 11368 11467: gap of 100 bp
* 11468 88955: contig of 77488 bp in length
* 88956 89055: gap of 100 bp
* 89056 155765: contig of 66710 bp in length
* 155766 155865: gap of 100 bp
* 155866 180726: contig of 24861 bp in length.
*
* Location/Qualifiers
* 1. 180726
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="11"
* /map="11"
* /clone_lib="RP11-784A24"
* /clone_id="RP11-784A24"
* /note="assembly_fragment"
* clone_end:SP6
* vector_side:left
* 4658_5364
* /note="assembly_fragment"
* 5465_7283
* /note="assembly_fragment"
* 7384_11367
* /note="assembly_fragment"
* 11468_88955
* /note="assembly_fragment"
* 89056_155765
* /note="assembly_fragment"
* 155866_180726
* /note="assembly_fragment"
* clone_end:T7
* vector_side:right
*
* ORIGIN
*
* Query Match 97.4%; Score 645; DB 2; Length 180726;
* Best Local Similarity 100.0%; Pred. No. 1.7e-142;
* Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
*
* QY 18 CTCTCTGCACGCGCCGCCGCCACCCACCATGCGCCACAGTTCAGCAGTGAAGGAAG 77
```

```
|||||
22603 CTCTCTGCAGCCGCGCCGCGCACCCACCCAGTTCAGAGCTGGAAGGAA 22544
QY 78 ATGGCGCTGTGGGACAGAAAGGCTTTGATGAATACATGAAGGAGCTAGAGTGGGAAT 137
Db 22543 ATGGCGCTGTGGGACAGAAAGGCTTTGATGAATACATGAAGGAGCTAGAGTGGGAAT 22484
QY 138 AGCTTTGCGAAAAATGGCGCAATGGCAAGCCAGATGTATCATCATCTGTGATGGTAA 197
Db 22483 AGCTTTGCGAAAAATGGCGCAATGGCAAGCCAGATGTATCATCATCTGTGATGGTAA 22424
QY 198 AAACCTCCACCATAAAACTGAGAGCACCTTTGAAAAACAACACAGTCTTTCTGTACCTCGG 257
Db 22423 AAACCTCCACCATAAAACTGAGAGCACCTTTGAAAAACAACACAGTCTTTCTGTACCTCGG 22364
QY 258 AGAGAAGTTTGAAGAAACACACAGCTGATGGCGAGAAAAAATCATGCTGTCTGCACTTTAC 317
Db 22363 AGAGAAGTTTGAAGAAACACACAGCTGATGGCGAGAAAAAATCATGCTGTCTGCACTTTAC 22304
QY 318 AGATGGTGCATTTGGTTACGATCAGAGTGGGATGGGAGGAAGCAACAATAACAAGAAA 377
Db 22303 AGATGGTGCATTTGGTTACGATCAGAGTGGGATGGGAGGAAGCAACAATAACAAGAAA 22244
QY 378 ATTGAAGATGGGAATTTAGTGGTGGAGTGTCTATGAACAATGTCACTGTCTGCTCGGAT 437
Db 22243 ATTGAAGATGGGAATTTAGTGGTGGAGTGTCTATGAACAATGTCACTGTCTGCTCGGAT 22184
QY 438 CTATGAAAAAGTGAATAAAAAATTCATCATCATCTTGGACAGGAGTTAAATTAAGAGAAT 497
Db 22183 CTATGAAAAAGTGAATAAAAAATTCATCATCATCTTGGACAGGAGTTAAATTAAGAGAAT 22124
QY 498 GACCAAGCTCAGTTCAATGAGCAATCTCCATCTGTTCTTTCTTTTTCATTATAC 557
Db 22123 GACCAAGCTCAGTTCAATGAGCAATCTCCATCTGTTCTTTCTTTTTCATTATAC 22064
QY 558 TGTGTTCAATTTATCTTTATATAACATTTTACATGCAGCTATTTTCAAAGTGTGTGGAT 617
Db 22063 TGTGTTCAATTTATCTTTATATAACATTTTACATGCAGCTATTTTCAAAGTGTGTGGAT 22004
QY 618 TAATTAGGATCATCCCTTTGGTTAATAATAATGTTGTGCT 662
Db 22003 TAATTAGGATCATCCCTTTGGTTAATAATAATGTTGTGCT 21959

RESULT 14
AC021809
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-568F15 map 11, WORKING DRAFT
SEQUENCE, 4 unordered pieces.
ACCESSION
AC021809.5 GI:14030015
VERSION
HTG; HTGS PHASE1; HTGS DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 195396)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE
Homo sapiens chromosome 11, clone RP11-568F15
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 195396)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Boukngalter,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Collangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,
Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
```

```

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 13, 2001 this sequence version replaced gi:12061487.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: I5111
Center clone name: 568_F15
-----
Summary Statistics
Sequencing vector: M13; M7815; 45% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 194127 bases at least Q40
Consensus quality: 194644 bases at least Q30
Consensus quality: 194869 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 195096; sum-of-fragments
Quality coverage: 11.4 in Q20 bases; agarose-fp
Quality coverage: 11.1 in Q20.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 48151: contig of 48151 bp in length
* 48152 48251: gap of 100 bp
* 48252 126200: contig of 77949 bp in length
* 126201 126300: gap of 100 bp
* 126301 162491: contig of 36191 bp in length
* 162492 162591: gap of 100 bp
* 162592 195396: contig of 32805 bp in length.
* 195396: Location/Qualifiers
* 1..195396
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="11"
* /map="11"
* /clone="RP11-568F15"
* /clone_lib="RPC1-11 Human Male BAC"
* 1..48151
* /note="assembly_fragment"
* clone_end:896
* vector_side:left"
* 48252..126200
* /note="assembly_fragment"
* 126301..162491
* /note="assembly_fragment"
* 162592..195396
* /note="assembly_fragment"
* clone_end:17
* vector_side:right"
misc_feature
1..48151
clone_end:896
vector_side:left"
misc_feature
48252..126200
note="assembly_fragment"
126301..162491
note="assembly_fragment"
162592..195396
note="assembly_fragment"
clone_end:17
vector_side:right"
ORIGIN
Query Match 97.4%; Score 645; DB 2; Length 195396;
Best Local Similarity 100.0%; Pred. No. 1.7e-142; Indels 0;
Matches 645; Conservative 0; Mismatches 0; Gaps 0;
```

```
QY 18 CTCTCTGCAGCGCCGCCCGCCGACCCACCATGCGCCACAGTTTCAGCAGCTGGAAGAAAG 77
Db 85685 CTCTCTGCAGCGCCGCCCGCCGACCCACCATGCGCCACAGTTTCAGCAGCTGGAAGAAAG 85744
QY 78 ATGGCGCCTGGTGGACAGCAAAAGGCTTTCATGAATACATGAAGGAGCTAGGAGTGGGAAAT 137
Db 85745 ATGGCGCCTGGTGGACAGCAAAAGGCTTTCATGAATACATGAAGGAGCTAGGAGTGGGAAAT 85804
QY 138 AGCTTTTGGCAAAATATGGCGGCAATGGCCAAAGCCAGATTTGTATCATCATCTTGTGATGGTAA 197
Db 85805 AGCTTTTGGCAAAATATGGCGGCAATGGCCAAAGCCAGATTTGTATCATCATCTTGTGATGGTAA 85864
QY 198 AAACCTCACATAAAATCTGAGAGCACTTTGAAAACAAACACAGTTTCTTGTACCTGGG 257
Db 85865 AAACCTCACATAAAATCTGAGAGCACTTTGAAAACAAACACAGTTTCTTGTACCTGGG 85924
QY 258 AGAGAAGTTTGAAGAAACACACAGCTGTATGCGAGAAAACCTCAGACTGTCTGCAACTTTTAC 317
Db 85925 AGAGAAGTTTGAAGAAACACACAGCTGTATGCGAGAAAACCTCAGACTGTCTGCAACTTTTAC 85984
QY 318 AGATGTGTGATTTGGTTTTCAGCATCAGAGTGGGATGGGAAGGAAAGCAACAATAACAAGAAA 377
Db 85985 AGATGTGTGATTTGGTTTTCAGCATCAGAGTGGGATGGGAAGGAAAGCAACAATAACAAGAAA 86044
QY 378 ATTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTTCATGAACAATGTCACTGTACTCGGAT 437
Db 86045 ATTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTTCATGAACAATGTCACTGTACTCGGAT 86104
QY 438 CTATGAAAAGTAGAATAAAATTCATCATCTTGGACAGGAGTGAATTAAGAGAT 497
Db 86105 CTATGAAAAGTAGAATAAAATTCATCATCTTGGACAGGAGTGAATTAAGAGAT 86164
QY 498 GACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTCTTTTCTTTTCTTTTCTTTT 557
Db 86165 GACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTCTTTTCTTTTCTTTTCTTTT 86224
QY 558 TGTGTTCAATATCTTTTATCATAAACATTTTATCATGCACTATTTCAAAAGTGTGTTGGAT 617
Db 86225 TGTGTTCAATATCTTTTATCATAAACATTTTATCATGCACTATTTCAAAAGTGTGTTGGAT 86284
QY 618 TAATTAGGATCATCCCTTGGTTTAAATAAATGTTTGTGCT 662
Db 86285 TAATTAGGATCATCCCTTGGTTTAAATAAATGTTTGTGCT 86329
```

RESULT 15

CQ725017 661 bp DNA linear PAT 03-FEB-2004
Sequence 10951 from Patent WO02068579.

ACCESSION CQ725017

VERSION CQ725017.1 GI:42285874

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

1

Venter C.J., Adame M.C., Li P.W. and Myers E.W.

Kit, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses

thereof

JOURNAL Patent: WO 02068579-A 10951 06-SEP-2002;

PE Corporation (NY) (US)

FEATURES

source

1. .861

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 97.0%; Score 642.4; DB 6; Length 661;

Best Local Similarity 98.3%; Pred. No. 7.8e-142;

```
Matches 649; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 3 CGCCGACGCGAGCCCTCTCTGCAGCGCCCGCCGACCCACCATGCGCCACAGTTTCAGTTCA 62
Db 2 CGCCGACGCGAGCCCTCTCTGCATGCGCCCGCCGACCCACCATGCGCCACAGTTCA 61
QY 63 GCAGCTGGAAGGAATGCGCGCTGCTGTCAGACGAAAGGCTTTGATGATACATGAGGA 122
Db 62 GCAGCTGGAAGGAATGCGCGCTGCTGTCAGACGAAAGGCTTTGATGATACATGAGGA 121
QY 123 GCTAGGAGTGGGAATAGCTTTTCCGAAAAATGCGCGCAATGCGCAAGCCAGGATTTGATCAT 182
Db 122 GCTAGGAGTGGGAATAGCTTTTCCGAAAAATGCGCGCAATGCGCAAGCCAGGATTTGATCAT 181
QY 183 CACTTGTGATGTTAAAAAACCCTCACCATAAAAACTGAGAGCACTTTGAAAAACAACAGTT 242
Db 182 CACTTGTGATGTTAAAAAACCCTCACCATAAAAACTGAGAGCACTTTGAAAAACAACAGTT 241
QY 243 TTTCTTTACCTCGGGAGAGAGTTTTGAAGAAACCAAGCTGATGCGAGAAAACCTCAGAC 302
Db 242 TTTCTTTACCTCGGGAGAGAGTTTTGAAGAAACCAAGCTGATGCGAGAAAACCTCAGAC 301
QY 303 TGTCTGCAACTTTTACAGATGGTGCATTTGTTTCCAGCATCAGGAGTGGGATGGGAAGAAAG 362
Db 302 TGTCTGCAACTTTTACAGATGGTGCATTTGTTTCCAGCATCAGGAGTGGGATGGGAAGAAAG 361
QY 363 CACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTGATGAACATGT 422
Db 362 CACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTGATGAACATGT 421
QY 423 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCTTTTGGACAGGA 482
Db 422 CACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTTCCATCATCTTTTGGACAGGA 481
QY 483 GTTAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTCT 542
Db 482 GTTAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTCT 541
QY 543 TTTTCTTTTCAATGTTTCAATTTATCTTTTATCATAAACATTTTACATGAGCTATTTT 602
Db 542 TTTTCTTTTCAATGTTTCAATTTATCTTTTATCATAAACATTTTACATGAGCTATTTT 601
QY 603 CAAAGTGTGTTGGATTAATTAGGATCATCCCTTTTGGTTAATAAATAAATGTTTGTGCT 662
Db 602 CAAAGTGTGTTGGATTAATTAGGATCATCCCTTTTGGTTAATAAATAAATGTTTGTGCT 661
```

Search completed: July 12, 2005, 17:41:17
Job time : 3060.23 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:38 ; Search time 251.715 Seconds
(without alignments)
8314.170 Million cell updates/sec

Title: US-09-788-074-2
Perfect score: 1275
Sequence: 1 aatgggagcaacatgctagc.....tctatgagaagtgcattga 1279

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:
2: /cgn2_6/prodata/1/ina/5B COMB.seq:
3: /cgn2_6/prodata/1/ina/6A COMB.seq:
4: /cgn2_6/prodata/1/ina/6B COMB.seq:
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:
6: /cgn2_6/prodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	9.8	408	1	US-08-446-600A-3
2	125	9.8	45762	4	US-09-949-016-16651
3	76.4	6.0	8324	4	US-09-949-016-13282
4	73.4	5.7	354	3	US-09-643-597-143
5	73.4	5.7	354	4	US-09-480-884A-143
6	73.4	5.7	354	4	US-09-542-615A-143
7	73.4	5.7	354	4	US-09-606-421B-143
8	73.4	5.7	354	4	US-09-221-107-143
9	73.4	5.7	354	4	US-09-466-396A-143
10	73.4	5.7	354	4	US-09-476-496A-143
11	73.4	5.7	354	4	US-09-630-940B-143
12	73.4	5.7	354	4	US-09-285-479-143
13	70	5.5	620	4	US-09-949-016-1540
14	65.6	5.1	601	4	US-09-949-016-42044
15	65.6	5.1	8597	4	US-09-949-016-12945
16	61.2	4.8	9739	4	US-09-949-016-12131
17	61.2	4.8	9739	4	US-09-949-016-13281
18	60.8	4.8	463	4	US-09-513-999C-14348
19	59.8	4.7	731	3	US-09-043-646-1
20	59.8	4.7	731	4	US-09-971-187-1
21	59.8	4.7	1022	4	US-09-949-016-1203
22	56.6	4.4	7218	1	US-08-232-463-14
23	56.2	4.4	493	4	US-09-513-999C-3851
24	56.2	4.4	606	4	US-09-513-999C-3961
25	56.2	4.4	2149	4	US-09-949-016-1539
26	56.2	4.4	2150	4	US-09-949-016-389
27	55.2	4.3	429	4	US-09-799-451-387

28	54.6	4.3	611	4	US-09-621-976-12879	Sequence 12879, A
29	50.6	4.0	505	4	US-09-621-976-15639	Sequence 15639, A
30	47.8	3.7	1717	1	US-08-468-709B-6	Sequence 6, Appli
31	47.8	3.7	1717	2	US-08-241-664B-6	Sequence 6, Appli
32	47.8	3.7	1717	5	PCT-US93-03936-6	Sequence 6, Appli
33	44.6	3.5	384	3	US-09-404-879A-333	Sequence 333, App
34	44.6	3.5	384	4	US-09-667-857-333	Sequence 333, App
35	44.6	3.5	466	4	US-09-702-705-61	Sequence 61, Appl
36	44.6	3.5	466	4	US-09-736-457-61	Sequence 61, Appl
37	44.6	3.5	466	4	US-09-614-124B-61	Sequence 61, Appl
38	44.6	3.5	466	4	US-09-671-325-61	Sequence 61, Appl
39	44.6	3.5	466	4	US-09-589-184-61	Sequence 61, Appl
40	44.6	3.5	466	4	US-09-658-824-61	Sequence 61, Appl
41	44.6	3.5	924	1	US-08-468-709B-1	Sequence 1, Appli
42	44.6	3.5	924	2	US-08-241-664B-1	Sequence 1, Appli
43	44.6	3.5	924	4	US-09-640-173-174	Sequence 174, App
44	44.6	3.5	924	4	US-09-713-550-174	Sequence 174, App
45	44.6	3.5	924	4	US-09-825-294-174	Sequence 174, App

ALIGNMENTS

RESULT 1
US-08-446-600A-3

; Sequence 3, Application US/08446600A
; Patent No. 5719126

; GENERAL INFORMATION:

; APPLICANT: No. 5719126dlund, James J. and Farcoqui, Jamal Z.

; TITLE OF INVENTION: MELANOGENIC INHIBITOR, AND METHODS OF PRODUCING AND USING THE

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Frost & Jacobs

; STREET: 2500 PNC Center, 201 East Fifth St.

; CITY: Cincinnati

; STATE: OH

; COUNTRY: USA

; ZIP: 45202-4182

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

; FILING DATE: 24 May 1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/980,513

; FILING DATE: 24 No. 5719126member 1992

; APPLICATION NUMBER: PCT/US93/11139

; FILING DATE: 16 No. 5719126member 1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Ann G. Robinson

; REGISTRATION NUMBER: 39,820

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (513) 651-6128

; TELEFAX: (513) 651-6981

; TELEX: 21-4396 F&J Cln

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 408 bases

; TYPE: nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: Unknown

US-08-446-600A-3

Query Match 9.8%; Score 125; DB 1; Length 408;

Best Local Similarity 78.8%; Pred. No. 4.1e-30;

Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189

Db 78 AGAGTGGGAATAGCTTTGCGAAAAATGGCCCATATGGCCCAAGCCAGATGTGTATCATCAC 137
QY 190 GTGTATGCGCAACAACATCACCGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
Db 138 TTGTATGTTAAACACCTCACCATAAACACTGAGAGCACTTTGAAAAACAACAGCTTTTC 197
QY 250 TTGTAACTGGGAGAGAGTGTGATGAACGACAGCTGATGCGAGAAAACTGAGGTGAG 309
Db 198 TTGTACCTGGGAGAGAGTGTGGAAGAAACCAACGCTGATGCGAGAAAACTCAGACTGT 257
QY 310 CTACAACAT 318
Db 258 CTGCAACTT 266

RESULT 2
US-09-949-016-16651/c
; Sequence 16651, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16651
; LENGTH: 45762
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16651

Query Match
Best Local Similarity 9.8%; Score 125; DB 4; Length 45762;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 130 AGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCCAAGCCAGACTGTATCATAC 189
Db 28038 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCCAAGCCAGACTGTATCATCAC 27979
QY 190 GTGTATGCGCAACAACATCACCGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
Db 27978 TTGTATGTTAAACACCTCACCATAAACCTGAGAGCACTTTGAAAAACAACAGTTTTTC 27919
QY 250 TTGTAACTGGGAGAGAGTGTGATGAACCGAGCTGATGCGAGAAAACTGAGGTGAG 309
Db 27918 TTGTACCTGGGAGAGAGTGTGGAAGAAACCAACGCTGATGCGAGAAAACTCAGACTGT 27859
QY 310 CTACAACAT 318
Db 27858 CTGCAACTT 27850

RESULT 3
US-09-949-016-13282
; Sequence 13282, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13282
; LENGTH: 8524
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13282

Query Match
Best Local Similarity 6.0%; Score 76.4; DB 4; Length 8524;
Matches 125; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 117 TTTCATATACCCACAGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCCAAGCCAG 176
Db 4602 TTTTCTTAAGTGTAGGAGTGGCTTTGCCACCAAGGAAGTGGCTGGCATGGCCCAACCTA 4661
QY 177 ACTGTATCATTCAGTGTGATGCGCAACAAACATCACCGTCAAAACCGAGAGCAGCTGAGA 236
Db 4662 ACATGATCATCAGTGTGAATGGGATGGATGCATCACCATTTAAATCTGAAAGTACCTTTAAA 4721
QY 237 CGACTGTCTTCTTGTAACTGGGAGAGAGTGTGATGAACGAGCAGCTGATGCGAGAA 296
Db 4722 ATACTGAGATTTCCTTCATCTAGGCGCCAGGAATTTGACGAAGTCACTGCAGATGACAGA 4781
QY 297 AAATCTGAGTCTAGCTACCAATACATCTG 322
Db 4782 AAGTCAAGGTGAGAAATAGGNAAGT 4807

RESULT 4
US-09-643-597-143
; Sequence 143, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-143

Query Match
Best Local Similarity 5.7%; Score 73.4; DB 3; Length 354;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 630 CAGACGGTCTGCACCTTCCAGACGGTGCCTGCTCCAGACCAGCAATGGGACGGGAAG 689
Db 22 CAGACTGTCTGCAACTTTACAGATGGTGCATTTGTTTACAGATCAGAGTGGGATGGGAAG 81
QY 690 GAGAGCAGGATAAACAAGAAACTGAAGGATGGGAAGATGATCGTGGTGAGCATCA 744

Db 82 GAAAGCACATACACAGAAAATTGAAAGATGGGAAATAGTGGTGGAGTGTGTCA 136

RESULT 5

US-09-480-884A-143
; Sequence 143, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-480-884A-143

Query Match 5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGCGGTGCCCTGTCAGACCAAGCAATGGGACGGGAAG 689

Db 22 CAGACTGTCTGCAACTTTACAGATGGTGCATTGTTTCAGCATCAGGAGTGGGATGGGAAG 81

QY 690 GAGAGCAGCATACACAGAACTGAGGATGGGAGATGATCTGTGACATCA 744

Db 82 GAAAGCACATACACAGAAAATTGAAAGATGGGAAATAGTGGTGGAGTGTGTCA 136

RESULT 6

US-09-542-615A-143
; Sequence 143, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-143

Query Match 5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGCGGTGCCCTGTCAGACCAAGCAATGGGACGGGAAG 689

Db 22 CAGACTGTCTGCAACTTTACAGATGGTGCATTGTTTCAGCATCAGGAGTGGGATGGGAAG 81

QY 690 GAGAGCAGCATACACAGAACTGAAAGATGGGAGATGATCTGTGACATCA 744

Db 82 GAAAGCACATACACAGAAAATTGAAAGATGGGAAATAGTGGTGGAGTGTGTCA 136

RESULT 7

US-09-606-421B-143
; Sequence 143, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-606-421B-143

Query Match 5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGCGGTGCCCTGTCAGACCAAGCAATGGGACGGGAAG 689

Db 22 CAGACTGTCTGCAACTTTACAGATGGTGCATTGTTTCAGCATCAGGAGTGGGATGGGAAG 81

QY 690 GAGAGCAGCATACACAGAACTGAAAGATGGGAGATGATCTGTGACATCA 744

Db 82 GAAAGCACATACACAGAAAATTGAAAGATGGGAAATAGTGGTGGAGTGTGTCA 136

RESULT 8

US-09-221-107-143
; Sequence 143, Application US/09221107
; Patent No. 6660838
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C2
; CURRENT APPLICATION NUMBER: US/09/221,107
; CURRENT FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Human
US-09-221-107-143

Query Match 5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGCGGTGCCCTGTCAGACCAAGCAATGGGACGGGAAG 689

Db 22 CAGACTGTCTGCAACTTTACAGATGGTGCATTGTTTCAGCATCAGGAGTGGGATGGGAAG 81

QY 690 GAGAGCAGCATACACAGAACTGAAAGATGGGAGATGATCTGTGACATCA 744

Db 82 GAAAGCACATACACAGAAAATTGAAAGATGGGAAATAGTGGTGGAGTGTGTCA 136

```
RESULT 9
US-09-466-396A-143
; Sequence 143, Application US/09466396A
; Patent No. 6696247
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-143

Query Match          5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGACGGTGCCCTGGTCCAGCACCAGCAATGGGACGGGAAG 689
Db 22 CAGACTGCTGCAACTTTACAGATGGTGCAATTGGTTTCAGCATCAGAGATGGGATGGGAAG 81

QY 690 GAGAGCAGCATTAACAAGAAAACCTGAAGGATGGGAAGATGATCGTGGTGAGCATCA 744
Db 82 GAAAGCACAAATAACAAGAAAATTGAAAGATGGGAAAATTAGTGGTGGAGTGTGTCA 136

RESULT 10
US-09-476-496A-143
; Sequence 143, Application US/09476496A
; Patent No. 6706262
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: LUNG CANCER
; FILE REFERENCE: 210121.455C5
; CURRENT APPLICATION NUMBER: US/09/476,496A
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-476-496A-143

Query Match          5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGACGGTGCCCTGGTCCAGCACCAGCAATGGGACGGGAAG 689
Db 22 CAGACTGCTGCAACTTTACAGATGGTGCAATTGGTTTCAGCATCAGAGATGGGATGGGAAG 81

QY 690 GAGAGCAGCATTAACAAGAAAACCTGAAGGATGGGAAGATGATCGTGGTGAGCATCA 744
Db 82 GAAAGCACAAATAACAAGAAAATTGAAAGATGGGAAAATTAGTGGTGGAGTGTGTCA 136

RESULT 11
US-09-630-940B-143
; Sequence 143, Application US/09630940B
; Patent No. 6737514
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C10
; CURRENT APPLICATION NUMBER: US/09/630,940B
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-630-940B-143

Query Match          5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGACGGTGCCCTGGTCCAGCACCAGCAATGGGACGGGAAG 689
Db 22 CAGACTGCTGCAACTTTACAGATGGTGCAATTGGTTTCAGCATCAGAGATGGGATGGGAAG 81

QY 690 GAGAGCAGCATTAACAAGAAAACCTGAAGGATGGGAAGATGATCGTGGTGAGCATCA 744
Db 82 GAAAGCACAAATAACAAGAAAATTGAAAGATGGGAAAATTAGTGGTGGAGTGTGTCA 136

RESULT 12
US-09-285-479-143
; Sequence 143, Application US/09285479
; Patent No. 6821518
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: LUNG CANCER
; FILE REFERENCE: 210121.455C3
; CURRENT APPLICATION NUMBER: US/09/285,479
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-285-479-143

Query Match          5.7%; Score 73.4; DB 4; Length 354;
Best Local Similarity 77.4%; Pred. No. 2.7e-13;
Matches 89; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 630 CAGACGGTCTGCACCTTCCAAAGACGGTGCCCTGGTCCAGCACCAGCAATGGGACGGGAAG 689
Db 22 CAGACTGCTGCAACTTTACAGATGGTGCAATTGGTTTCAGCATCAGAGATGGGATGGGAAG 81

QY 690 GAGAGCAGCATTAACAAGAAAACCTGAAGGATGGGAAGATGATCGTGGTGAGCATCA 744
Db 82 GAAAGCACAAATAACAAGAAAATTGAAAGATGGGAAAATTAGTGGTGGAGTGTGTCA 136

RESULT 13
US-09-949-016-1540
; Sequence 1540, Application US/09949016
```


Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1540
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1540

Query Match 5.1%; Score 70; DB 4; Length 620;
Best Local Similarity 60.5%; Pred. No. 5.2e-12; Indels 0; Gaps 0;
Matches 115; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
DB 119 AGGAGTGGGCTTTGCCACAGGAAGTGGCTGCCATGGCCAACTTAACATGATCATCAG 178
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTTCTC 249
DB 179 TGTGAATGGGGATGTGATCACCATTAAATCTGAAGTACCTTTAAATACTAGAGATTTC 238
QY 250 TTGTACTCTGGGAGAGAACTTTGATCAACAGCAGCTGATGGCAGAAAACCTGAGTCTAG 309
DB 239 CTCATCTAGGGCCAGGAATTTGACGAAGTCACTGCAGATGACAGGAAGTCAAGAGCAC 298
QY 310 CTACAACATA 319
DB 299 CATAACCTTA 308

RESULT 14
US-09-949-016-42044
; Sequence 42044, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42044
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-42044

Query Match 5.1%; Score 65.6; DB 4; Length 601;
Best Local Similarity 58.0%; Pred. No. 1.4e-10;
Matches 116; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189

DB 362 AGCGTGGGCTTTGCCACTAGGCAGGTGGGAAATGTGACCAACCAACCGTAATTATCAG 421
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTTCTC 249
DB 422 TCAGAGAGGAGACAAAGTGGTCAATCAGGACTCTCAGCACATTCAGAGACGAGATTAG 481
QY 250 TTGTAACTCTGGGAGAGAACTTTGATGAAAACGACAGCTGATGGCAGAAAACCTGAGTCTAG 309
DB 482 TTTCCAGCTGGGAGAGAGTTTGTATGAAACCACTGCAGATGATAGAAACTGTGAAGGTGAG 541
QY 310 CTACAACATACTGTGAAGCG 329
DB 542 AAACCTGCTTCTTCTTCAGAG 561
RESULT 15
US-09-949-016-12945
; Sequence 12945, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12945
; LENGTH: 8597
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12945

Query Match 5.1%; Score 65.6; DB 4; Length 8597;
Best Local Similarity 58.0%; Pred. No. 8.8e-10;
Matches 116; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
DB 2815 AGCGTGGGCTTTGCCACTAGGCAGGTGGGAAATGTGACCAACCAACCGTAATTATCAG 2874
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTTCTC 249
DB 2875 TCAGAGAGGAGACAAAGTGGTCAATCAGGACTCTCAGCACATTCAGAGACGAGATTAG 2934
QY 250 TTGTAACTCTGGGAGAGAACTTTGATGAAAACGACAGCTGATGGCAGAAAACCTGAGTCTAG 309
DB 2935 TTTCCAGCTGGGAGAGAGTTTGTATGAAACCACTGCAGATGATAGAAACTGTGAAGGTGAG 2994
QY 310 CTACAACATACTGTGAAGCG 329
DB 2995 AAACCTGCTTCTTCTTCAGAG 3014

Search completed: July 12, 2005, 18:08:06
Job time : 253.715 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:34 ; Search time 807.859 Seconds
(without alignments)
9372.119 Million cell updates/sec

Title: US-09-788-074-2
Perfect score: 1279
Sequence: 1 aatgggagcaacatgctagc.....tctatgagaagtgcaatga 1279

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1279	100.0	1279	4	AAS13246 Mouse DNA
2	284	22.2	3247	9	ACC42875 Adenosine
3	176.6	13.8	933	12	ADJ75985 Marker ge
4	175.6	13.7	537	12	ACH72029 Human gen
5	165	12.9	1623	6	ABS65621 Mouse Fab
6	144.2	11.3	643	6	ABK71797 Human dit
7	144.2	11.3	695	10	ADF30639 Rat angio
8	144.2	11.3	704	10	ADF30635 Rat angio
9	144.2	11.3	704	12	ADP72791 Renal tox
10	142.6	11.1	664	10	ADB53766 Primary r
11	142.6	11.1	664	10	ADP30637 Rat angio
12	129.8	10.1	335	6	ABL92950 Rat metas
13	129.8	10.1	335	12	ADN07741 Human mam
14	125	9.8	408	2	AAQ66842 Melanogen
15	125	9.8	461	6	ABK53980 Human hea
16	125	9.8	479	9	ACH38376 Human end
17	125	9.8	494	9	ACH35892 Human end
18	125	9.8	519	6	ABK53994 Human hea
19	125	9.8	606	6	ABK53952 Human hea
20	125	9.8	612	10	ADK11762 Breast ca

21	125	9.8	615	6	ABK53828 Human hea
22	125	9.8	662	4	AAS13247 Human DNA
23	125	9.8	662	6	ABV77978 Hypoxia-r
24	125	9.8	662	10	ADH75278 Prostate
25	125	9.8	662	10	ADH28827 Human chr
26	125	9.8	662	12	ADJ75120 Marker ge
27	125	9.8	662	12	ADN03852 Antipsoori
28	125	9.8	662	12	ADO19263 Human PRO
29	125	9.8	662	12	ADP13321 Renal cel
30	125	9.8	662	13	ADR24747 Breast ca
31	125	9.8	662	13	ACN38822 Tumour-as
32	125	9.8	662	13	ADP54337 Human PRO
33	125	9.8	662	13	ADR52981 Drug ther
34	125	9.8	662	13	ADP25373 PRO polyo
35	125	9.8	662	13	ADR99018 Fatty aci
36	125	9.8	662	13	ADR66234 Human pro
37	125	9.8	662	13	ADR66576 Human pro
38	125	9.8	720	2	AAZ77538 Human ova
39	125	9.8	1071	10	ADB47413 Human CDN
40	125	9.8	1071	11	ADM86791 Human DNA
41	125	9.8	1072	6	AAS94888 Human DNA
42	124	9.7	331	6	ABK53860 Human hea
43	123.4	9.6	627	9	AAD57420 Human FAB
44	121.8	9.5	540	9	AAD57419 Human fat
45	121.8	9.5	660	12	ADQ86530 Human tum

ALIGNMENTS

RESULT 1
AAS13246
ID AAS13246 standard; DNA; 1279 BP.
XX
AC AAS13246;
XX

DT 18-DEC-2001 (first entry)
XX

DE Mouse DNA encoding keratinocyte fatty acid binding protein, Mall.
KW Mouse; Mal 1; keratinocyte fatty acid binding protein; ds; anorectic;
KW antidiabetic; antilipemic; antiatherosclerotic; obesity; diabetes;
KW dyslipidaemia; atherosclerosis; antisense therapy.
XX

OS Mus musculus.

FH Key Location/Qualifiers
FT CDS 132..1279

FT exon /product= "Mal 1"

FT /tag= a

FT /tag= b

FT /number= 2

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

FT /number= 1

XX 10-APR-2001; 2001CN-00105888.
XX PF
XX PR
XX 10-APR-2001; 2001CN-00105888.
XX (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.
XX PA
XX PI Mao Y, Xie Y;
XX WPI; 2003-230977/23.
XX P-PSDB; ABR56128.
XX A polypeptide-adenosine triphosphate synthetase-10.89, encoding
PT polynucleotide, antagonist, and recombinant production, useful for
PT treating hypertension, peptic ulcers, nephrotic syndrome, asthma and
PT parkinsonism.
XX PS
XX Claim 6; Page 26-28; 31pp; Chinese.
XX CC
XX The present invention relates to adenosine triphosphate synthetase-10.89
CC and its coding sequence. The protein is useful for treating hypertension,
CC peptic ulcers, nephrotic syndrome, asthma and parkinsonism
XX SQ
XX Sequence 3247 BP; 1007 A; 549 C; 683 G; 1008 T; 0 U; 0 Other;
Query Match 22.2%; Score 284; DB 9; Length 3247;
Best Local Similarity 59.2%; Pred. No. 5.7e-77;
Matches 720; Conservative 0; Mismatches 425; Indels 71; Gaps 11;
QY 126 CCACAGGAGTAGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCA 185
DB 1839 CTACAGGATGGGAATAGCTTTGCGAAATAATGGCGCAATGGCCAGCCAGATTGTATCA 1898
QY 186 TTACGTGTGATGCAACAAACATCACCGTCAAAACCGAGAGCAGTGAAGACGACTGTGT 245
DB 1899 TCATTTGTGATGTAATAAACCTCACCATATAAAGCTGAGAGCATTGTGAATAACACACAGT 1958
QY 246 TCTTTGTAACTGGGAGAGAAGTTTGTGATAACGACAGCTGATGGCAGAAAAAAGTGGG 305
DB 1959 TTTCTTGT-CCCTGGGAGAGAAGTTTGAAGAAACACACAGCTGATGGCAGAAAAAAGTGGG 2017
QY 306 TCAGCTACACATACATGTCGAGGACAGAGCTCTAGATTACAGATTAAATTCATTA 365
DB 2018 TCAGTCGTGACATGTTATGAATACAGAAAGCTCTAGAAATGATAGGCTGTAT---CAAT 2074
QY 366 ACAATGCTGTACTTACTGCGAAGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTT 425
DB 2075 AACATTTTACTGTTTATAGGCAAGAACTTATGAATAAAGATTATTTAT-GAATTGAATTT 2133
QY 426 TGATAAATTAGTAAAGTCCCAGGACTAAGAAATGAAGACATCTTATGAGTTTCTAGATC 485
DB 2134 TGTCAAATTAGCAAAAGTATCAACTTCATCATAGAATTGGCATCTTTTATTAGCTACTAG 2193
QY 486 GAAAGCACATAGTTGTATGTGAACAAATCAGTATGATGGGTGGAGT-----TCAGAG 541
DB 2194 GTTGAACACCAAACTATTTGTAATAAATAATCAATATGGTTTAAATGAAGTAGACTCAGAA 2253
QY 542 AGGGAAGCGGAAGACTTGTGTGAGT-----GGTGGGTCTGGGGGTCTCTTCACTT 595
DB 2254 AGGAGAGGTGAACAAATGTTGATTGAAGAGTTATGATCATGGAATACTCTTGTATG 2313
QY 596 TGAAGATGATGAATCACTACCTGTATTTTTTGCAGACGGTGTGCACCTTCCAGACGG 655
DB 2314 TACTTGAAGATTAAACCGTTTACTTTGTTTTTGCAGACTGTCTGCAACTTTTACAGATGG 2373
QY 656 TGCCTTGTTCAGCACCAAGCAATGGGACGGGAAGGAGACGATGAACAGAAACTGAA 715
DB 2374 TGCATTGGTTTCAGCATCAGAGTGGGATGGGAAGGAAGCAAAATGAAGAAATTTGAA 2433
QY 716 GATGGGAAGATGATCGTGTGAGCATCAAGCAGCTGGCACCCTGCTGGGATTTGGSCCTG 775
DB 2434 AGATGGAAATTAGTGGTGAAGTGTCAACTGCTGTCTCAGTCAGCTTCTTGTGTG 2493

QY 776 CAGCCACAGTTGTCTATACCACTTCGGGTCACTTGGTGTCTTTTAAACAAGAGAAACTTAA 835
DB 2494 CATTCATAGTTCACATAAAGTCTTCTATATCATGATCAIT-----AACAGAACTCAGTTT 2549
QY 836 GGAGGACATACCTGAAAATAACAAGTTAGAAACGAGAGTCCCTCATTGCTGAGCAGCCCT 895
DB 2550 GAAAGAAAAAAGCAAAATAACAAGTTAAATCTAGAACACTAATTTATTAAAGAAAGTCT 2609
QY 896 TGTGGGACGGAAGTGTATGGG-----ATCCCAGGATGTGGCTGACAGCAGAGCCTG---- 947
DB 2610 AGTGGAGATAGAGAAGTGCATGACATGACATGAGAGAGGATTTGGCTGGAGTTGGGGGGAGTTC 2669
QY 948 -----AGAGCTGCAGGCCACCGAGCAGCCCTCTCTCTGGT 982
DB 2670 TTGCTTTGTGTCACGTCACAGTGAATCTTGGCAAGCCCAAACTGCACACATTTTTC 2729
QY 983 ACATTTGATTTAAGTAAGGATATTTGCCAAAAACACATGAATTAATTTAGAGATCATATCCA 1042
DB 2730 TATATCTTTAAATAAGGAATATTTGCCATAACCAAAAGAAATAAGATATATTCAACACA 2789
QY 1043 GTGCTTTAGTCTGCGAGGCGCAAAATATATACATATATAACAAAAACAGCAGCTCTAGGTCCT- 1101
DB 2790 TTAGTTTTCTGTAATAACATGCATATTTATAAGCAAAAAACACAGCTTCTGGCTTCTCTCAA 2849
QY 1102 CTTGAGTTTGAATCTGAGATGTGTTTTTCTGTAGGTGGTTTACAGGGTTTATAGGA 1161
DB 2850 CCGGTGAATTTCTGAAAGAAATTTTCTCCATCTATGAAGTAGATTACGTGATTTTCGTGGGA 2909
QY 1162 TTCTGCC-----ACAAACACATGCTCTGAAATGTACAGTTGCGCCTGAGACTC 1208
DB 2910 CTTTGATTTCTAATGTTTAAATACCACTGCTCTGGAATCTAAGGCTAACCTAACTCTT 2969
QY 1209 T-----ATCTTTCTTCTCTAGGAGTGTCTATGAACAATGCCACCTGCACCTCGGCTCTA 1263
DB 2970 TTAATATCTTCTTCTCTAGGAGTGTCTCATGAACAATGTCACCTGTACTCGGATCTA 3029
QY 1264 TGAGAAGGTGCAATGA 1279
DB 3030 TGAATAAGTAGAATAA 3045
RESULT 3
ADJ75985
ID ADJ75985 standard; DNA; 933 BP.
XX AC
XX ADJ75985;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene SEQ ID NO:1237.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker gene; gene; ds.
XX OS Mus musculus.
XX PN EP1394274-A2.
XX PD 03-MAR-2004.
XX PF 04-AUG-2003; 2003EP-00254857.
XX PR 06-AUG-2002; 2002JP-00229312.
XX PR 20-MAR-2003; 2003JP-00077212.
XX PA (GENO-) GENOX RES INC.
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuha K;
XX WPI; 2004-193155/19.
XX DR
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by

PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.

XX Example 11; SEQ ID NO 1237; 241pp; English.

XX The present invention describes a method of testing for bronchial asthma
XX or chronic obstructive pulmonary disease. The method comprises
XX determining the expression level of a marker gene in a biological sample
XX from a subject, comparing the expression level determined with the
XX expression level of the marker gene in a biological sample from a healthy
XX subject, and judging whether the subject has bronchial asthma or chronic
XX obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX genes (S1) whose expression levels increase when respiratory epithelial
XX cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX whose expression levels decrease when respiratory epithelial cells are
XX stimulated with interleukin-13. Also described: (1) a reagent (1) for
XX testing for bronchial asthma or chronic obstructive pulmonary disease;
XX (2) a kit for screening for a candidate compound for a therapeutic agent
XX to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX an animal model for bronchial asthma or chronic obstructive pulmonary
XX disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX method for producing an animal model for bronchial asthma or chronic
XX obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX asthma or chronic obstructive pulmonary disease, comprising the compound,
XX a marker gene or an antisense nucleic acid corresponding to a portion of
XX the marker gene, a ribozyme, a polynucleotide that suppresses the
XX expression of the gene through an RNAi effect or an antibody recognising
XX a protein encoded by a marker gene; and (7) a DNA chip for testing for
XX bronchial asthma or a chronic obstructive pulmonary disease, on which a
XX probe has been immobilised to assay a marker gene. (1) has respiratory
XX and antiasthmatic activities, and can be used in gene therapy. The method
XX is useful for testing for or screening for a therapeutic agent for
XX bronchial asthma or chronic obstructive pulmonary disease. The present
XX sequence is used in the exemplification of the present invention.

SQ Sequence 933 BP; 263 A; 188 C; 228 G; 254 T; 0 U; 0 Other;

Query Match 13.8%; Score 176.6; DB 12; Length 933;
Best Local Similarity 95.3%; Pred. No. 8.8e-44;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGGATGAGGACCTGCTTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATAC 189
DB 112 AGAGTAGGACCTGCTTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATAC 171
QY 190 GTGTGATGGCAACATCATCCGTCATCAAAACCGAGAGCAGCAGTGAAGACGCTGTCTTC 249
DB 172 GTGTGATGGCAACATCATCCGTCATCAAAACCGAGAGCAGCAGTGAAGACGCTGTCTTC 231
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGTATGGCAGAAAACCTGAGTCAG 309
DB 232 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGTATGGCAGAAAACCTGAGTCAG 291
QY 310 CTACACATAC 320
DB 292 CTGACCTTCC 302

RESULT 4

ID ACH72029 standard; DNA; 537 BP.

XX ACH72029;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #5224.

XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.

XX Homo sapiens.

XX

PN US2003194704-A1.

PD 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

PI Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

DR New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.

XX Claim 15; SEQ ID NO 5224; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
XX sequences in the specification, or their complements or fragments, and
XX encoding at least 8 amino acids of any of the 888 amino acid sequences
XX fully defined in the specification. The probe is a single exon probe that
XX hybridises under high stringency conditions to a nucleic acid molecule
XX expressed in human cells or tissues. Also included are a spatially-
XX addressable set of single exon nucleic acid probes for measuring human
XX gene expression (comprising a plurality of single exon nucleic acid
XX probes cited above, where each of the plurality of probes is separately
XX and addressably isolatable or amplifiable from the plurality), a single
XX exon microarray for measuring human gene expression, a method of
XX measuring human gene expression, a vector comprising the single exon
XX probe cited above, an ORF-encoded peptide comprising at least 8
XX contiguous amino acids of any of the above-mentioned amino acid
XX sequences (optionally with conservative amino acid substitutions), an
XX isolated antibody that binds specifically to a peptide cited above,
XX methods of selling and/or licensing single exon probes or microarrays to
XX a customer desiring to measure gene expression, a method of providing
XX human gene expression data by subexpression, and a computer-readable
XX storage medium which contains a database having a plurality of records
XX (each record including data on the expression of a single exon probe
XX cited above. The probe, methods and apparatus are useful in gene
XX expression analysis. The probes may be used as tools for surveying
XX tissues to detect the presence of expressed messages that contain their
XX specific exon, or in constructing genome-derived single exon microarrays.
XX In addition, the probes are used in identifying and characterising
XX alternative splicing events, in detecting and characterising gross
XX alterations in the genomic locus that includes their exon, in assessing
XX smaller genomic alterations, in priming the synthesis of nucleic acids,
XX or in expressing the ORF-encoded peptide. The present sequence is a human
XX single exon probe of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from USPTO at
XX seqdata.uspto.gov/sequence.html?DocID=20030194704

XX Sequence 537 BP; 197 A; 80 C; 104 G; 156 T; 0 U; 0 Other;

Query Match 13.7%; Score 175.6; DB 12; Length 537;
Best Local Similarity 67.6%; Pred. No. 1.3e-43;
Matches 277; Conservative 0; Mismatches 129; Indels 4; Gaps 2;

QY 126 CCACAGGAGTAGGACTGGCTTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCA 185
DB 98 CTACAGGAGTGGGATAGCTTTGCGAAAATGGCGCAATGGCCAGCAGCTGTATCA 157
QY 186 TTACGTGTGATGGCAACATCATCCGTCATCAAAACCGAGAGCAGCAGTGAAGACCTGTCT 245
DB 158 TCACCTGTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCAGCTTTGAAAACACACAGT 217

QY 246 TCTCTTGAACCTGGGAGAGAGCTTTGATGAACAGCAGCTGATGGCAGAAAACTGAGG 305
DB 218 TTTCTTTGATCCCTGGGAGAGAGCTTTGAAAGAAACACAGCTGATGGCAGAAAACTCAGG 277
QY 306 TCAGCTACAACTACTGTGAAGCGACAGAGAGCTTCTAGATTATACAGATTAAATTGCATTA 365
DB 278 TCAGTCGTGACATGTTATGAATATCAGAGAGCTTCTAGAAATGATAGCTGTAT---CAAT 334
QY 366 ACAATCTCTGTACTTACTGCTCAGAGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTT 425
DB 335 AACATTTTACTGTTTATAGCAAGAACTTAATGAAAAAGTTATTTAT-GAATTGAATTT 393
QY 426 TGAATAATTAGTAAAGTCCAGGACTAAGAAATGAAGACATCTTATGAGTTTCTAGATC 485
DB 394 TGTCAAATTAGCAAAAGTATCACTTCATCATAGAAATGGCATCTTTTATAGTACTACTAG 453
QY 486 GAAAGCACATAGTCTGTTATGTAACAAATCAGTATGATGGGTGAGT 535
DB 454 GTTGAAGAACACAACTATTGTGAATAAATCAATATGGTTAATGAAGT 503

RESULT 5
ABS65621
ID ABS65621 standard; DNA; 1623 BP.
XX AC ABS65621;
XX DT 15-NOV-2002 (first entry)
XX DE Mouse Fabe gene (EMBL No. fabe_mouse) exon 1.
XX KW Mouse; vertebrate; anorexia/anorexia genotype; anx/anx;
KW control of food intake; feed; body weight; human; weight gain;
KW ANKT protein; nucleolar protein; proliferating cell; anorectic;
KW chromosome 2; gene; ds.
XX OS Mus musculus.
XX PN WO200262835-A1.
XX PD 15-AUG-2002.
XX PF 08-FEB-2002; 2002NO-SE000222.
XX PR 08-FEB-2001; 2001US-00778844.
XX PA (APPE-) APPETITE CONTROL AB.
XX PI Johansen J, Schalling M;
XX WPI; 2002-657520/70.

XX New nucleic acid molecule encoding a polypeptide capable of at least
XX partially complementing the phenotype of an animal having the anx/anx
XX genotype, useful for controlling and regulating weight gain and food
XX intake in humans.
XX Disclosure; Fig 11; 314pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
XX comprising a nucleotide sequence derived from a vertebrate animal (e.g.
XX mouse), and encodes a polypeptide capable of at least partially
XX complementing the phenotype of an animal having the anorexia/anorexia
XX (anx/anx) genotype. The nucleic acid molecule and the polypeptide encoded
XX by the nucleotide sequence is useful for controlling food or feed intake,
XX or body weight in animals including humans. The polynucleotide and
XX polypeptide sequences of the invention are useful for regulating and
XX controlling weight gain and food intake in animals particularly humans.
XX The polypeptide sequences are also useful for generating monoclonal or
XX polyclonal antibodies. The antibodies are useful for diagnostic and
XX therapeutic purposes. These antibodies can be used in an immunoassay for
XX identification or detection of the polypeptides. The polynucleotide

CC sequences are useful for recombinant production of large quantities of
CC the receptor, for further studies into the mechanism of regulation of
CC food intake and/or body weight and for obtaining agonists and
CC antagonists. ABS65557-ABS65762 represent mouse gene sequences relating to
CC the present invention. The gene sequences are located on mouse chromosome
CC 2
XX
SQ Sequence 1623 BP; 575 A; 332 C; 371 G; 345 T; 0 U; 0 Other;
Query Match 12.9%; Score 165; DB 6; Length 1623;
Best Local Similarity 92.1%; Pred. No. 5e-40;
Matches 174; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 130 AGAGTAGAGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
DB 612 AGAGTAGAGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 671
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCCAGAGACAGTGAACAGCACTGTGTCTC 249
DB 672 TTGTGATGGCAACAACATCACCGTCAAAACCCAGAGACAGTGAACAGCACTGTGTCTC 731
QY 250 TTGTAACTCTGGGAGAGAGCTTTGATGAAACGACAGCTGATGGCAGAAAACTGAGTCTAG 309
DB 732 TTGTACCTCTGGGAGAGAGCTTTGATGAAACGACAGCTGATGGCAGAAAACTGAGCGGT 791
QY 310 CTACAAACAT 318
DB 792 CTGCACCTT 800
RESULT 6
ABK71797
ID ABK71797 standard; cDNA; 643 BP.
XX AC ABK71797;
XX DT 30-JUL-2002 (first entry)
XX DE Human dithp polynucleotide #263.
XX KW Human; dithp; diagnostic and therapeutic polynucleotide; gene; ss; bone;
KW cell proliferative disorder; cancer; tumour; autoimmune disorder; brain;
KW inflammatory disorder; viral infection; bacterial infection; seizure;
KW fungal infection; parasitic infections; developmental disorder; breast;
KW endocrine disorder; metabolic disorder; neurological disorder; cervix;
KW gastrointestinal disorder; transport disorder; gene therapy; kidney;
KW adrenal gland; bone marrow; lung; ovary; pancreas; prostate; spleen;
KW skin; testis; thymus.
XX OS Homo sapiens.
XX PN WO200220754-A2.
XX PD 14-MAR-2002.
XX PF 29-AUG-2001; 2001WO-US027127.
XX PR 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 06-SEP-2000; 2000US-0230583P.
PR 06-SEP-2000; 2000US-0230505P.
PR 06-SEP-2000; 2000US-0230514P.
PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.
PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.
PR 06-SEP-2000; 2000US-0230595P.
PR 06-SEP-2000; 2000US-0230597P.
PR 06-SEP-2000; 2000US-0230598P.
PR 06-SEP-2000; 2000US-0230599P.

```
PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230865P.
PR 06-SEP-2000; 2000US-0230988P.
PR 07-SEP-2000; 2000US-0230951P.
PR 07-SEP-2000; 2000US-0231163P.
PR 07-SEP-2000; 2000US-0231167P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;
PI Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
PI Momiyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
PI Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
PI Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;
XX
XX WPI; 2002-383054/41.
DR P-PSDB; ABG60208.
XX
XX An isolated polynucleotide useful in diagnostics and therapeutics.
XX
XX Claim 1; Page 525; 686pp; English.
XX
XX The invention relates to human diagnostic and therapeutic (dithp)
CC polynucleotides and their associated polypeptides (DITHP polypeptides).
CC The sequences of the invention are used in the treatment and diagnosis of
CC cell proliferative disorders (e.g. atherosclerosis, cirrhosis), cancers
CC (e.g. tumours of the adrenal gland, bone, bone marrow, brain, breast,
CC cervix, kidney, lung, ovary, pancreas, prostate, skin, spleen, testis or
CC thymus), autoimmune/inflammatory disorders (e.g. asthma, bronchitis,
CC psoriasis, osteoporosis), viral infections, bacterial infections, fungal
CC infections, parasitic infections, developmental disorders (e.g. anaemia,
CC epilepsy), seizure disorders (e.g. cerebral palsy, spina bifida),
CC endocrine disorders (e.g. thrombosis, aneurysm), metabolic disorders
CC (e.g. obesity, diabetes), neurological disorders (e.g. stroke,
CC amyotrophic lateral sclerosis, multiple sclerosis), gastrointestinal
CC disorders (e.g. ulcerative colitis, lysinuria) and transport disorders
CC (e.g. myotonic dystrophy, catatonia, peripheral neuropathy). Sequences
CC ABK71535-ABK71809 represent human dithp polynucleotides of the invention
XX
XX Sequence 643 BP; 173 A; 146 C; 177 G; 147 T; 0 U; 0 Other;
SQ
Query Match 11.3%; Score 144.2; DB 6; Length 643;
Best Local Similarity 85.2%; Pred. No. 9.1e-34;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
Db 95 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 154
QY 190 GTGTGATGGCAACACATCACCGTCAAAACCGAGACAGTGAAGACGACTGTGTCTC 249
Db 155 CCTGACGGCAACAACCTCACCGTCAAAACCTGAGAGACGGTGAAGACCGCTGTGTTC 214
QY 250 TTGTAACTTGGGAGAGAACTTTGATGAACACGACAGCTGATGGCAGAAAACCTGAGTCCAG 309
Db 215 TTGCACCTTGGGAGAGAACTTTGATGAACACGACAGCTGATGGCAGAAAACCTGAGACGCT 274
QY 310 CTACAACAT 318
Db 275 CTGCACCTT 283
RESULT 7
ADF30639
ID ADF30639 standard; cDNA; 695 BP.
XX
XX ADF30639;
XX
XX 12-FEB-2004 (first entry)
DT
DE Rat angiogenesis modulating protein cDNA #53.
XX
XX ss; gene; rat; angiogenesis; angiogenesis modulating protein;
KW
```

```
KW retinal neovascularisation; choroidal neovascularisation;
KW chronic inflammation; myocardial ischaemia; stroke;
KW coronary artery disease; peripheral vascular disease.
XX
XX Rattus norvegicus.
XX
XX US2003162706-A1.
XX
XX 28-AUG-2003.
XX
XX 10-DEC-2002; 2002US-00316253.
XX
XX 08-FEB-2002; 2002US-0355295P.
XX 26-JUN-2002; 2002US-0391758P.
XX
XX (PROC ) PROCTER & GAMBLE CO.
XX
XX Peters KG, Thompson LJ, Wang F, Greis KD;
XX
XX WPI; 2003-711557/67.
DR P-PSDB; ADF30640.
XX
XX Treating angiogenesis-mediated disorder, e.g., retinal or choroidal
PT neovascularization or diseases associated with chronic inflammation,
PT myocardial ischemia, stroke, coronary artery disease or peripheral
PT vascular disease.
XX
XX Disclosure; SEQ ID NO 202; 26pp; English.
XX
XX The invention relates to a method of treating an angiogenesis-mediated
CC disorder in a subject. The method is useful for treating angiogenesis-
CC mediated disorder, e.g., retinal or choroidal neovascularisation or
CC diseases associated with chronic inflammation, myocardial ischaemia,
CC stroke, coronary artery disease or peripheral vascular disease. The
CC present sequence is used in the exemplification of the invention.
XX
XX Sequence 695 BP; 194 A; 162 C; 183 G; 156 T; 0 U; 0 Other;
SQ
Query Match 11.3%; Score 144.2; DB 10; Length 695;
Best Local Similarity 85.2%; Pred. No. 9.5e-34;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
Db 136 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 195
QY 190 GTGTGATGGCAACACATCACCGTCAAAACCGAGACAGTGAAGACGACTGTGTCTC 249
Db 196 CCTGACGGCAACAACCTCACCGTCAAAACCTGAGAGACGGTGAAGACCGCTGTGTTC 255
QY 250 TTGTAACTTGGGAGAGAACTTTGATGAACACGACAGCTGATGGCAGAAAACCTGAGTCCAG 309
Db 256 TTGCACCTTGGGAGAGAACTTTGATGAACACGACAGCTGATGGCAGAAAACCTGAGACGCT 315
QY 310 CTACAACAT 318
Db 316 CTGCACCTT 324
RESULT 8
ADF30635
ID ADF30635 standard; cDNA; 704 BP.
XX
XX ADF30635;
XX
XX 12-FEB-2004 (first entry)
DT
DE Rat angiogenesis modulating protein cDNA #51.
XX
XX ss; gene; rat; angiogenesis; angiogenesis modulating protein;
KW retinal neovascularisation; choroidal neovascularisation;
KW chronic inflammation; myocardial ischaemia; stroke;
KW coronary artery disease; peripheral vascular disease.
```



```
XX OS Rattus norvegicus.
XX PN US2003162706-A1.
XX PD 28-AUG-2003.
XX PF 10-DEC-2002; 2002US-00316253.
XX PR 08-FEB-2002; 2002US-0355295P.
XX PR 26-JUN-2002; 2002US-0391758P.
XX PA (PROC ) PROCTER & GAMBLE CO.
XX PI Peters KG, Thompson LJ, Wang F, Greis KD;
XX WPI; 2003-711557/67.
XX DR P-PSDB; ADF30636.
XX PT Treating angiogenesis-mediated disorder, e.g., retinal or choroidal
PT neovascularization or diseases associated with chronic inflammation,
PT myocardial ischemia, stroke, coronary artery disease or peripheral
PT vascular disease.
XX PS Disclosure; SEQ ID NO 198; 26pp; English.
XX CC The invention relates to a method of treating an angiogenesis-mediated
CC disorder in a subject. The method is useful for treating angiogenesis-
CC mediated disorder, e.g., retinal or choroidal neovascularisation or
CC diseases associated with chronic inflammation, myocardial ischaemia,
CC stroke, coronary artery disease or peripheral vascular disease. The
CC present sequence is used in the exemplification of the invention.
XX SQ Sequence 704 BP; 217 A; 157 C; 178 G; 152 T; 0 U; 0 Other;
Query Match 11.3%; Score 144.2; DB 10; Length 704;
Best Local Similarity 85.2%; Pred. No. 9.5e-34;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGTCCATGGCCAGCCAGACTGTATCATTTAC 189
DB 121 AGGAGTAGGCTGGCTCTTAGGAGATGGTGGCCATGGCCAAACAGATGTCATCATTTAC 180
QY 190 GTGTGTATGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAAGACGACTGTGTCTC 249
DB 181 CCTGACGGCAACAACCTCACCGTCAAAACTGAGAGACCGGTGAAGACGACCGTGTTC 240
QY 250 TTGTAACTCTGGAGAGAGATTTGATCAAAACGAGACAGTGATGGCAAAAACCTGAGTCAG 309
DB 241 TTGCACCTTTGGGAGAGAGATTTGATGAAACCAACAGCTGATGGCAGGAAAACCTGAGACG 300
QY 310 CTACAAACAT 318
DB 301 CTGCACCTT 309
RESULT 9
ID ADP72791 standard; DNA; 704 BP.
XX AC ADP72791;
XX DT 26-AUG-2004 (first entry)
XX DE Renal toxin progression gene marker #1380.
XX KW ds; toxic effect; gene expression profile; kidney tissue;
XX KW differential gene expression; toxicity progression; toxicity marker;
XX KW drug screening; toxicity assay; kidney pathology; nephritis;
XX KW kidney necrosis; glomerular injury; tubular injury;
XX KW focal segmental glomerulosclerosis.
XX OS Rattus norvegicus.
```

```
XX WO2004048598-A2.
XX PD 10-JUN-2004.
XX PF 24-NOV-2003; 2003WO-US037556.
XX PR 22-NOV-2002; 2002US-00301856.
XX PA (GENE-) GENE LOGIC INC.
XX PI Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;
XX PI Elashoff M;
XX WPI; 2004-460771/43.
XX PT Predicting (the progression of) a toxic effect of a compound, for
XX monitoring the progression of renal disease states, comprises preparing a
XX gene expression profile of a kidney tissue or cell sample exposed to the
XX compound.
XX PS Claim 11; SEQ ID NO 1380; 266pp; English.
XX CC The invention relates to a method of predicting (the progression of) a
XX toxic effect of a compound by preparing a gene expression profile of a
XX kidney tissue or cell sample exposed to the compound and comparing the
XX gene expression profile to a database, or detecting the level of gene(s)
XX expression in a tissue or cell sample exposed to the compound, where
XX differential gene expression compared to a control indicates a toxic
XX effect (toxicity progression). The method is useful for predicting (the
XX progression of) at least one toxic effect of a compound. The genes are
XX useful as toxicity markers in drug screening and toxicity assays. The
XX methods are useful for predicting the likelihood that a compound or test
XX agent will induce various specific kidney pathologies, such as nephritis,
XX kidney necrosis, glomerular and tubular injury, or focal segmental
XX glomerulosclerosis. The methods are useful for determining the similarity
XX of a toxic response to one or more individual compounds and for
XX predicting or elucidating the potential cellular pathways influenced,
XX induced or modulated by the compound or test agent. The kit is useful for
XX predicting or modelling the toxic response of a test compound, for
XX monitoring the progression of renal disease states, for identifying genes
XX that show promise as new drug targets and for screening known and newly
XX designed drugs. This sequence corresponds to a gene marker used in the
XX method of the invention. (Note: The sequence data for this patent did not
XX form part of the printed specification, but was obtained in electronic
XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 704 BP; 217 A; 157 C; 178 G; 152 T; 0 U; 0 Other;
Query Match 11.3%; Score 144.2; DB 12; Length 704;
Best Local Similarity 85.2%; Pred. No. 9.5e-34;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGTCCATGGCCAGCCAGACTGTATCATTTAC 189
DB 121 AGGAGTAGGCTGGCTCTTAGGAAGATGGTGGCCATGGCCAAACAGATGTCATCATTTAC 180
QY 190 GTGTGTATGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAAGACGACTGTGTCTC 249
DB 181 CCTGACGGCAACAACCTCACCGTCAAAACTGAGAGACCGGTGAAGACGACCGTGTTC 240
QY 250 TTGTAACTCTGGGAGAGAGATTTGATGAAACGAGCTGATGGCAGAAAACCTGAGTCAG 309
DB 241 TTGCACCTTTGGGAGAGAGATTTGATGAAACCAACAGCTGATGGCAGGAAAACCTGAGACG 300
QY 310 CTACAAACAT 318
DB 301 CTGCACCTT 309
RESULT 10
ID ADB53766 standard; DNA; 664 BP.
```

```
XX ADB53766;
AC
XX
XX
DT 04-DEC-2003 (first entry)
DE
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4308.
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; db.
XX
OS Rattus norvegicus.
XX
XX WO2003065993-A2.
XX
XX 14-AUG-2003.
XX
XX 04-FEB-2003; 2003WO-US003482.
XX
XX 04-FEB-2002; 2002US-0353171P.
XX 13-MAR-2002; 2002US-0363534P.
XX 08-APR-2002; 2002US-0370248P.
XX 10-APR-2002; 2002US-0371134P.
XX 10-APR-2002; 2002US-0371135P.
XX 10-APR-2002; 2002US-0371150P.
XX 11-APR-2002; 2002US-0371413P.
XX 19-APR-2002; 2002US-0373601P.
XX 19-APR-2002; 2002US-0373602P.
XX 22-APR-2002; 2002US-0374139P.
XX 08-MAY-2002; 2002US-0378370P.
XX 09-MAY-2002; 2002US-0378652P.
XX 09-MAY-2002; 2002US-0378653P.
XX 09-MAY-2002; 2002US-0378665P.
XX 09-JUL-2002; 2002US-0394230P.
XX 09-JUL-2002; 2002US-0394253P.
XX 04-SEP-2002; 2002US-0407688P.
XX 28-JAN-2003; 2003US-0442900P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
XX Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
XX identifying hepatotoxic compounds, comprises comparing a gene expression
XX profile of a tissue or cell sample to a database of Tox mean and non-Tox
XX mean values.
XX
XX Claim 44; SEQ ID NO 4308; 874pp; English.
XX
XX The present invention describes a method for determining whether a
XX compound induces a toxic effect on a tissue or cell. The method comprises
XX preparing a gene expression profile of a tissue or cell sample exposed to
XX the compound, and comparing the gene expression profile to a database
XX comprising data or information on the Tox mean and non-Tox mean value.
XX The method is useful for predicting or identifying at least one toxic
XX effect, particularly hepatotoxicity, of a test or unknown compound. The
XX genes listed in the specification are useful as diagnostic or toxicity
XX markers for the prediction or identification of the physiological state
XX of tissue or cell sample that has been exposed to a compound, or to
XX identify or predict the toxic effects of a compound or an agent. These
XX may also be used as markers for monitoring toxicity progression or for
XX drug screening. The present sequence represents a primary rat hepatocyte
XX toxicity modelling related gene sequence from the present invention.
XX
XX Sequence 664 BP; 180 A; 152 C; 178 G; 154 T; 0 U; 0 Other;
SQ
Query Match 11.1%; Score 142.6; DB 10; Length 664;
Best Local Similarity 84.7%; Pred. No. 2.9e-33;
Matches 160; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATTTAC 189
DB 119 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATTTAC 178
QY 190 GTGTGATGGCAACACATCAGGTCACAAACCGAGAGCAGCTGAAGACGACTGTCTTC 249
DB 179 CCTGCACACACACACCTCACCGTCACAACTGAGAGCAGCTGAAGACGACCGTGTTC 238
QY 250 TTGTAACCTGGGAGAGAGATTGATGAACACACAGCTGATGCGCAAAAACCTGAGCTCAG 309
DB 239 TTGCACCTTGGGAGAGAGATTGATGAACACACAGCTGATGCGCAAAAACCTGAGACCGT 298
QY 310 CTACAACAT 318
DB 299 CTGCACCTT 307
RESULT 11
ADF30637
ID ADF30637 standard; cDNA; 664 BP.
XX
AC ADF30637;
XX
DT 12-FEB-2004 (first entry)
XX
DE Rat angiogenesis modulating protein cDNA #52.
XX
KW ss; gene; rat; angiogenesis; angiogenesis modulating protein;
KW retinal neovascularisation; choroidal neovascularisation;
KW chronic inflammation; myocardial ischaemia; stroke;
KW coronary artery disease; peripheral vascular disease.
XX
OS Rattus norvegicus.
XX
XX US2003162706-A1.
XX
XX 28-AUG-2003.
XX
XX 10-DEC-2002; 2002US-00316253.
XX
XX 08-FEB-2002; 2002US-0355295P.
XX 26-JUN-2002; 2002US-0391758P.
XX
XX (PROC ) PROCTER & GAMBLE CO.
XX
XX Peters KG, Thompson LJ, Wang F, Greis KD;
XX
XX WPI; 2003-711557/67.
XX P-PSDB; ADF30638.
XX
XX Treating angiogenesis-mediated disorder, e.g., retinal or choroidal
XX neovascularization or diseases associated with chronic inflammation,
XX myocardial ischaemia, stroke, coronary artery disease or peripheral
XX vascular disease.
XX
XX Disclosure; SEQ ID NO 200; 26pp; English.
XX
XX The invention relates to a method of treating an angiogenesis-mediated
XX disorder in a subject. The method is useful for treating angiogenesis-
XX mediated disorder, e.g., retinal or choroidal neovascularisation or
XX diseases associated with chronic inflammation, myocardial ischaemia,
XX stroke, coronary artery disease or peripheral vascular disease. The
XX present sequence is used in the exemplification of the invention.
XX
XX Sequence 664 BP; 180 A; 152 C; 178 G; 154 T; 0 U; 0 Other;
SQ
Query Match 11.1%; Score 142.6; DB 10; Length 664;
Best Local Similarity 84.7%; Pred. No. 2.9e-33;
Matches 160; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATTTAC 189
DB 119 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCAGCTGTATCATTTAC 178
```


QY 130 AGGACTAGGACTGGCTCTTAGGAAGATGGCTGCC-ATGCCAAGCCAGACTGTATCATTA 188
|||||
Db 305 AGGAGTAGGGCTGGCTCTTAGGGAGATGGGTGCMATGGCCAAACCCAGACTGCATCATTA 246
|||||
QY 189 CGTGTGATGGCAACAACTACCGGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTCT 248
|||||
Db 245 CCCTAGACGGCNAACAACTCACCGTMAAACTGAGAGCACGGTGAAGAMGACCGTGT 186
|||||
QY 249 CTTGTAACTGGGAGAGAAGTTTGATGAAACGACAGCTGATGGCAGAAAAAATGAGGTCA 308
|||||
Db 185 CTTGCACCTTGGGAGAGAAGTTTGATGAAACCCACAGCTGATGGCAGGAAAAAATGAGACGG 126
|||||
QY 309 GCTACAACATA 319
|||||
Db 125 TCTGCACCTTA 115
|||||
RESULT 14
AAQ66842
ID AAQ66842 standard; DNA; 408 BP.
XX
AC AAQ66842;
XX
DT 25-MAR-2003 (revised)
DT 07-DEC-1994 (first entry)
XX
DE Melanogenic inhibitor.
XX
KW Melanogenic inhibitor; MI; melanogenesis; melanin; pigment;
KW depigmentation; melanoma; ss.
XX
OS Homo sapiens.
XX
FH Key
FT 1. 408
FT CDS /*tag= a
XX
XX WO9412534-A2.
PN
PD 09-JUN-1994. 93WO-US011139.
XX
PF 16-NOV-1993;
XX
PR 24-NOV-1992; 92US-00980513.
PR 01-SEP-1993; 93US-00115172.
XX
PA (UYCI-) UNIV CINCINNATI.
XX
PI Nordlund JJ, Farooqui JZ;
XX
XX WPI; 1994-200198/24.
DR P-PSDB; AAR55866.
XX
XX Prodn. of a protein for inhibiting melanogenesis - useful for treating
PT hyper-pigmentary diseases, destroying melanoma cells and for lightening
PT unwanted body hair.
XX
PS Disclosure; Page 3-4; 40pp; English.
XX
XX PCR primers given in AAQ66843-44 were used to amplify RNA from human skin
CC previously grafted onto nude mice, thereby providing DNA (AAQ66842)
CC encoding melanogenic inhibitor protein (AAR55866). (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 408 BP; 142 A; 67 C; 111 G; 88 T; 0 U; 0 Other;
Query Match 9.8%; Score 125; DB 2; Length 408;
Best Local Similarity 78.8%; Pred. No. 6.8e-28;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCAAGCCAGACTGTATCATTA 189
|||||
|||||

Db 78 AGGAGTGGAAATAGCTTTTCGAAAAAATGGCGCAATGGCCCAAGCCAGATTGTATCATCAC 137
QY 190 GTGTGATGGCAACAACTACCGGTCAAAACCGAGAGCACACTGAAGACGACTGTCTTCTC 249
|||||
Db 138 TTGTGATGGTAAAAACCTCACCAATAAAACCTTGAGAGCACTTTTGAANAACACACAGTTTTTC 197
|||||
QY 250 TTGTAACTCTGGGAGAGAAGTTTGATGAAACACACAGCTGATGGCAGAAAAAATCGAGGTGAG 309
|||||
Db 198 TTGTACCTCTGGGAGAGAAGTTTGAAGAAACCCACAGCTGATGGCAGAAAAAATCGAGACTGT 257
|||||
QY 310 CTACAACAT 318
|||||
Db 258 CTGCAACTT 266
|||||
RESULT 15
ABK53980
ID ABK53980 standard; cDNA; 461 BP.
XX
AC ABK53980;
XX
DT 05-JUN-2002 (first entry)
XX
DE Human head and neck tumour cDNA, SEQ ID No 175.
XX
KW Human; head and neck cancer; tumour; cytostatic; immunogenic; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212329-A2.
XX
PD 14-FEB-2002.
XX
PF 01-AUG-2001; 2001WO-US024226.
XX
PR 03-AUG-2000; 2000US-0223281P.
PR 16-NOV-2000; 2000US-0249933P.
XX
XX (CORI-) CORIXA CORP.
XX
PI Wang T, Fan L;
XX
XX WPI; 2002-257467/30.
DR
XX Novel polynucleotide encoding head and neck tumor polypeptides, useful in
PT pharmaceutical compositions, e.g. vaccines, for treating head and neck
PT cancers.
XX
PS Claim 1; Page 167; 200pp; English.
XX
CC The invention relates to an isolated polynucleotide (I) comprising
CC sequences selected from 273 sequences fully defined in the specification.
CC (I), including its encoded polypeptide (II), an antibody binding to (II),
CC a fusion protein comprising (II) and a T-cell population stimulated by
CC (I) or (II) are useful for stimulating an immune response in a patient
CC and treating head and neck cancer in a patient. An oligonucleotide (III)
CC that hybridises to (I) is useful for determining the presence of cancer
CC in a patient, by obtaining a biological sample from the patient,
CC contacting the sample with (III), detecting in the sample an amount of a
CC polynucleotide that hybridises to the oligonucleotide, and comparing the
CC amount of polynucleotide that hybridises to the oligonucleotides to a
CC predetermined cut-off value. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis and
CC treatment of head and neck cancer. ABK53806-ABK54078 represent human head
CC and neck cancer cDNA sequences of the invention
XX
SQ Sequence 461 BP; 153 A; 91 C; 119 G; 96 T; 0 U; 2 Other;
Query Match 9.8%; Score 125; DB 6; Length 461;
Best Local Similarity 78.8%; Pred. No. 7.3e-28;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

This Page Blank (uspto)


```
QY 181 TATCATTTAGTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCACAGTGAAGCAGC 240
Db 181 TATCATTTAGTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCACAGTGAAGCAGC 240

QY 241 TGTGTTCTCTTGTAACTCGGAGAGAGTGTGATGAAACGACAGCTGATGGCAGAAAAC 300
Db 241 TGTGTTCTCTTGTAACTCGGAGAGAGTGTGATGAAACGACAGCTGATGGCAGAAAAC 300

QY 301 TGAGGTGAGTCAACATCTGTGAAGCGACAGAAGCTTCTAGATTTACAGATTAATATG 360
Db 301 TGAGGTGAGTCAACATCTGTGAAGCGACAGAAGCTTCTAGATTTACAGATTAATATG 360

QY 361 CATTAACAATCTGTACTTACTGCGAAGGGCTGACTGAAAAAATCTTATATGAGTTG 420
Db 361 CATTAACAATCTGTACTTACTGCGAAGGGCTGACTGAAAAAATCTTATATGAGTTG 420

QY 421 ACTTTTGATAAATTAGTAAAGTCCAGGAGTAAAGAAATGAACACATCTTATGAGTTCT 480
Db 421 ACTTTTGATAAATTAGTAAAGTCCAGGAGTAAAGAAATGAACACATCTTATGAGTTCT 480

QY 481 AGATCGAAAGCACAATAGTTGTTATTTGTGAACAAATCAGTATGATGGGTGAGTTTCA 540
Db 481 AGATCGAAAGCACAATAGTTGTTATTTGTGAACAAATCAGTATGATGGGTGAGTTTCA 540

QY 541 GAGGGAAGCGGAGACTTGTGTGGAGTGTGTGGGTCTTGGGGTTCCTTCACTTTGGAA 600
Db 541 GAGGGAAGCGGAGACTTGTGTGGAGTGTGTGGGTCTTGGGGTTCCTTCACTTTGGAA 600

QY 601 GATGATGAATCTACTACCTGTTATTTTGCAGACGGTCTGCACCTTCCAAGCGGTGCC 660
Db 601 GATGATGAATCTACTACCTGTTATTTTGCAGACGGTCTGCACCTTCCAAGCGGTGCC 660

QY 661 TGCTCCAGCACCAGCAATGGGAGGAGAGACGATGATGATGATGATGATGATGATGATG 720
Db 661 TGCTCCAGCACCAGCAATGGGAGGAGAGACGATGATGATGATGATGATGATGATGATG 720

QY 721 GGAAGATGATCGTGTGAGCATCAAGACATGCGCACCATGCTGGGATGCGGCTGAGCC 780
Db 721 GGAAGATGATCGTGTGAGCATCAAGACATGCGCACCATGCTGGGATGCGGCTGAGCC 780

QY 781 ACAGTTGTCTATACCACTTCCGGTCAATGCTTTTAAAGAGAGAGGAACTTAGGAGG 840
Db 781 ACAGTTGTCTATACCACTTCCGGTCAATGCTTTTAAAGAGAGAGGAACTTAGGAGG 840

QY 841 ACAATCTGAAATAACAAGTTAGAAAACGAGAGTCTCTATGCTGAGGAGCCCTTTGTG 900
Db 841 ACAATCTGAAATAACAAGTTAGAAAACGAGAGTCTCTATGCTGAGGAGCCCTTTGTG 900

QY 901 GGACGAGAGAGTATGGGATCCAGGATGTTGGCTGAGAGAGCCTGAGAGCTGGCAGGC 960
Db 901 GGACGAGAGAGTATGGGATCCAGGATGTTGGCTGAGAGAGCCTGAGAGCTGGCAGGC 960

QY 961 CACCGAGAGCCCTCTCTGTTACATTTAAGTAAGGATATTTGCCAAACACATG 1020
Db 961 CACCGAGAGCCCTCTCTGTTACATTTAAGTAAGGATATTTGCCAAACACATG 1020

QY 1021 AATAATTTAGAGATCATATCCAGTGTCTTAGTCTGAGGAGCAGCAATATACATATAAC 1080
Db 1021 AATAATTTAGAGATCATATCCAGTGTCTTAGTCTGAGGAGCAGCAATATACATATAAC 1080

QY 1081 AAAACAGAGCTTAGTCTTCTGATGTTGAAATCTGAGATGTTGTTTCTTCTGTTAGT 1140
Db 1081 AAAACAGAGCTTAGTCTTCTGATGTTGAAATCTGAGATGTTGTTTCTTCTGTTAGT 1140

QY 1141 TGGTTACAGCGTTTATAGATTTCTCCCAACACATCTCTGAAATCTACAGTTGGCC 1200
Db 1141 TGGTTACAGCGTTTATAGATTTCTCCCAACACATCTCTGAAATCTACAGTTGGCC 1200

QY 1201 TGAGACTCTATCTTTCTCTCTAGGAGTGTGTGATGAAACAATGCCACCTGCACTCGGT 1260
Db 1201 TGAGACTCTATCTTTCTCTCTAGGAGTGTGTGATGAAACAATGCCACCTGCACTCGGT 1260
```

```
QY 1261 CTATGAGAAGTGCATGA 1279
Db 1261 CTATGAGAAGTGCATGA 1279

RESULT 3
MMFABPE
LOCUS MMFABPE Mus musculus Fabpe gene. 6593 bp DNA linear ROD 27-JUL-1998
DEFINITION AJ223066
VERSION AJ223066.1 GI:3183984
KEYWORDS E-FABP; epidermal protein; Fabpe gene; fatty acid binding protein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Bleck,B., Hohoff,C., Binas,B., Rustow,B., Dixkens,C., Hameister,H.,
Borchers,T. and Spener,F.
TITLE Cloning and chromosomal localisation of the murine epidermal-type
fatty acid binding protein gene (Fabpe)
JOURNAL Gene 215 (1), 123-130 (1998)
MEDLINE 98332726
PubMed 9666100
REFERENCE 2
AUTHORS Spener,F.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1997) Spener F., Department of Biochemistry,
University of Muenster, Wilhelm-Klemm-Str. 2, D-48149, GERMANY
REMARK Revised by [3]
AUTHORS 3 (bases 1 to 6593)
TITLE Spener,F. and Bleck,B.
JOURNAL Direct Submission
REMARK Submitted (01-JUN-1998) Spener F., Department of Biochemistry,
University of Muenster, Wilhelm-Klemm-Str. 2, D-48149, GERMANY
COMMENT On Jun 4, 1998 this sequence version replaced gi:2760452.
FEATURES
source
1..8593
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="SvJ"
/db_xref="taxon:10090"
/chromosome="3"
/map="3A1-3"
/cell_line="RW4 cells"
2470..6491
/gene="Fabpe"
2470..2582
/gene="Fabpe"
/number=1
Join(2504)..2582,4800..4972,5300..5401,5948..6001)
/gene="Fabpe"
/note="other name of mRNA: mall"
/codon_start=1
/product="epidermal-type fatty acid binding protein
(E-FABP)"
/protein_id="CAA11069.1"
/db_xref="GI:2760453"
/db_xref="GOA:Q05816"
/db_xref="UniProt/Swiss-Prot:Q05816"
/translation="MASLKDLGKWLMSHGFEYMKELGVGLALRKNAAKAPDCI
ITCDGNNITVKTSTVFCNIGKDFETADGRKRTETCTFODGALVHQHMDG
KESTITRKLDKGMIVCEVMNNATCTRYEVQ"
2583..4799
/gene="Fabpe"
/number=1
4800..4972
/gene="Fabpe"
/number=2
4973..5299
/gene="Fabpe"
/number=2
5300..5401
intron
exon
intron
exon
```

		/gene="Fabpe"		
		/number=3		
intron		5402..5947		
		/gene="Fabpe"		
		/number=3		
exon		5948..6491		
		/gene="Fabpe"		
		/number=4		
ORIGIN				
	Query Match	86.5%;	Score 1106.2;	DB 10; Length 6593;
	Best Local Similarity	94.7%;	Pred. No. 0;	
	Matches 1263; Conservative	0;	Mismatches 9;	Indels 62; Gaps 9;
QY	1	AATGGAGCAACATGCTAGCTATGCGAGTGC	60	
Db	4675	AATGGAGCAACATGCTAGCTATGCGAGTGC	4730	
QY	61	GGCCAGTGGGATGATAAGGAATCAATCCTTGCTTATCATTTGACAAATTTACGTCAATTTTC	120	
Db	4731	GGCCAGTGGGATGATAAGGAATCAATCCTTGCTTATTAATTTGACAAATTTATGTCAATTT--	4788	
QY	121	CATACCCACAGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTG	180	
Db	4789	CCATCCCAACAGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTG	4848	
QY	181	TATCAATTAGTGTGATGGCAACAAACATCACGGTCAAAACCGAGAGCACAGTGAAGACGAC	240	
Db	4849	TATCAATTAGTGTGATGGCAACAAACATCACGGTCAAAACCGAGAGCACAGTGAAGACGAC	4908	
QY	241	TGTGTTCTCTTTGTAACCTCGGAGAGAGTTTTGATGAACGACAGCTGATGGCAGAAAAAC	300	
Db	4909	TGTGTTCTCTTTGTAACCTCGGAGAGAGTTTTGATGAACGACAGCTGATGGCAGAAAAAC	4968	
QY	301	TGAGGTGAGTCAACATCTGTGAAGCGACAGAAAGCTTCTAGATTTTACAGATTAATATG	360	
Db	4969	TGAGGTGAGTCAACATCTGTGAAGCGACAGAAAGCTTCTAGATTTTACAGATTAATATG	5028	
QY	361	CATTAAATGTCGTACTTACTGCCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTG	420	
Db	5029	CATTAAATGTCGTACTTACTGCCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTG	5088	
QY	421	ACTTTTGATAAATTAGTAAAGTCCCAAGGACTCAAGAAATGAAGACATCTTATGAGTTTCT	480	
Db	5089	ACTTTTGATAAATTAGTAAAGTCCCAAGGACTCAAGAAATGAAGACATCTTATGAGTTTCT	5148	
QY	481	AGATCGAAAGCACATAGTTGTATTGTGAACAAAAATCAGTATGATGGGTGGAGTTTCAGA	540	
Db	5149	AGATCGAAAGCACATAGTTGTATTGTGAACAAAAATCAGTATGATGGGTGGAGTTTCAGA	5208	
QY	541	GAGGGAAGCGGAGAGCTCTTGTGAGTGGTGTGGGTCTCTGGGGTTCCTTCACTTTGGAA	600	
Db	5209	GAGGGAAGCGGAGAGCTCTTGTGAGTGGTGTGGGTCTCTGGGGTTCCTTCACTTTGGAA	5267	
QY	601	GATGATGAATTAATACCTGTAATTTTTCAGACGGTCTGCACCTTCCAGACGGTGGCC	660	
Db	5268	GATGATGAATTAATACCTGTAATTTTTCAGACGGTCTGCACCTTCCAGACGGTGGCC	5327	
QY	661	TGTTCCAGCACGACCAATGGGACGGGAAGCAGAGCACGATAACAAGAAAACTGAAGGATG	720	
Db	5328	TGTTCCAGCACGACCAATGGGACGGGAAGGAGAGCACGATAACAAGAAAACTGAAGGATG	5387	
QY	721	GGAAGATGATCGTGGTGAGCATC-AAAGCATGGCAACATGCTGGGATT--GGGCTCGAG	778	
Db	5388	GGAAGATGATCGTGGTGAGCATCAAAAGCATCTGGCACCATGCTGGGATTGGGGCTCGAG	5447	
QY	779	CCACAGTTGTCATAACCACTTCGGGTCAATGGTTCATTTTAAACAAGAGAGGAAACTTAGGA	838	
Db	5448	CCACAGTTGTCATAACCACTTCGGGTCAATGGTTCATTTTAAACAAGAGAGGAAACTTAGGA	5507	
QY	839	GGACAATCTGAAAAATAACAAGTTAGAAAACGAGAGTCTCTATTGCTGAGGCGACCTTGT	898	
Db	5508	GGACAATCTGAAAAATAACAAGTTAGAAAACGAGAGTCTCTATTGCTGAGGCGACCTTGT	5567	

QY	899	GGGACCGAGAGTCAATGGGA-TCCAGGATGTGGCTCAGCAGAG-CCTGAGAGCTGGC	956		
DB	5568	GGGACCGAGAGTCAATGGGA-TCCAGGATGTGGCTCAGCAGAGCCTGAGAGCTGGC	5627		
QY	957	AGGCCACCGAGCAGCCCTCTCTGTCATATTGAATTAAGTAAGGATATTTGCCAAAAACA	1016		
DB	5628	AGGCCACCGAGCAGCCCTCTCTGTCATATTGAATTAAGTAAGGATATTTGCCAAAAACA	5687		
QY	1017	CATGAATAATTAGAGATCATATCCAGTGTCTTGTAGTCTGCAGGGCAGCAATATATACATAT	1076		
DB	5688	CATGAATAATTAGAGATCATATCCAGTGTCTTGTAGTCTGCAGGGCAGCAATATATACATAT	5747		
QY	1077	AAACAAAC-CAGCAGCTCTAGCTCTTC-----	1102		
DB	5748	AAACAAAC-CAGCAGCTCTAGCTCTTC-----	5807		
QY	1103	-----TTGAGTTTGAATCCTGAGATGTGGTTTCTTCTGTAGTTGGTT	1145		
DB	5808	TAGTTTGGTTTGAATCCTGAGATGTGGTTTCTTCTGTAGTTGGTT	5867		
QY	1146	ACAGCGTTTATAGGATTTCTGCCCAACACATGCTCTGAAATGTACAGTTGGCTGAGA	1205		
DB	5868	ACAGCGTTTATAGGATTTCTGCCCAACACATGCTCTGAAATGTACAGTTGGCTGAGA	5927		
QY	1206	CTCTATCTTCTTCTCTAGGAGTGTGTCATGAACAAATGCCACCTGCACTCGGGTCTATG	1265		
DB	5928	CTCTATCTTCTTCTCTAGGAGTGTGTCATGAACAAATGCCACCTGCACTCGGGTCTATG	5987		
QY	1266	AGAAGTGCATGA 1279			
DB	5988	AGAAGTGCATGA 6001			
RESULT 4					
AC097097					
LOCUS					
DEFINITION					
Rattus norvegicus clone CH230-189N20, WORKING DRAFT SEQUENCE, 9					
AC097097					
unordered pieces.					
AC097097.5					
HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.					
KEYWORDS					
Rattus norvegicus (Norway rat)					
SOURCE					
Rattus norvegicus					
ORGANISM					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;					
Rattus.					
REFERENCE					
AUTHORS					
1 (bases 1 to 234778)					
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,					
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,					
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,					
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,					
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,					
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,					
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,					
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,					
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,					
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,					
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,					
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,					
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,					
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,					
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,					
Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,					
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,					
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,					
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,					
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,					
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,					
Karpthy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,					
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,					
Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,					
Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,					

Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minje, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sibson, I., Sittler, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmari, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission
 Unpublished
 2 (bases 1 to 234778)
 Worley, K.C.
 Direct Submission
 Submitted (09-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 234778)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 10, 2003 this sequence version replaced gi:23321796.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GILZ
 Center clone name: CH230-189N20
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 214384 bases at least Q40
 Consensus quality: 217402 bases at least Q30
 Consensus quality: 219588 bases at least Q20
 Estimated insert size: 224004; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

```
Db 172035 GGGTGGTGTGGGTGTT-AAGGCTCTGCACCTTTGGAGATT-----ATAAATACCTTGCA 172087
Qy 624 TTTTTCGACAGCGTCTGCACCTTCAAGACGCTGCCCTGGTCCAGCACCAGCAATGGGAC 683
Db 172088 TTTTTCGACAGCGTCTGCACCTTCAAGACGCTGCCCTGGTCCAGCACCAGCAATGGGAA 172147
Qy 684 GGGAGGAGAGCAGCAGTAAACAGAACTGAAGGATGGGAAGATGATCGTGGTGAGCATC 743
Db 172148 GGGAAAGAAAGCAGCAGTAAACAGAACTGAAGGACGGGAAGATGATCGTGGTGAGTGT 172207
Qy 744 AAGACACTGGCACCCTGCTGGGATT-GGGCCTGCGAGCCACAGTGTGTCATAACCACTTCGG 802
Db 172208 GAAGTGT--TATCATGCTGGCATTTCTGGCTTGCAGCCACTTGTGCACA----- 172253
Qy 803 GTCATGTTCTTTTAAACAGAGAACTTGAAGGACATATCTGAAATTAACAAGTT 862
Db 172254 -----ACCATGAAGAGAGACTTAGGAAGAAAGTATTTCAGACGCGATGTG 172299
Qy 863 AGAAACGAGAGTCTCATTTGCTGAGCGAGCCCTTGTGGGACGGAGAAGTATGGGATCC 922
Db 172300 ATTCCAGAT-----GTGGCTGGAGCAGTGTCTTCC 172331
Qy 923 CAGGATGTGGCTGCGACGAGCCTGAGAGCTGGCAGGCCACCGAGCAGCCCTCTCTCGT 982
Db 172332 TGGGCTGTGGCAGTGTG-----TGAGAGCTGGCAGGCTATCAAGCAGCCCTCTCTCGT 172386
Qy 983 ACATTGATTT-----AAGTAAGGATATTGTCACAAA-CACATGAATAT 1026
Db 172387 ACATATGTTTTCCTCCCGTTTATCTAAAGGAAGGCTATTGTCACAAACCATAGTAT 172446
Qy 1027 TTAGAGATCATATCAGTCTTTAGTCTGCAGGCGAGCAAAATATACATATAACAAACA 1086
Db 172447 TTAGGATCATTTCCAGTCTCTAGTCTCCAGATAGCAAGCTTCGTATATAACAAACA 172506
Qy 1087 GCAGCTCTAGG---TCTTCTTGAGTTTGAATCCTGAGATGTGGTTTTCTGT-----AGG 1139
Db 172507 GCAGCTCTCAGTGTGTTTCCAGTCCAGAAATCCTGATGTGTCTTCTTCTGTAGCAAGG 172566
Qy 1140 TTGGTTACAGGCTTATAGGATTTCTGCCACACACATGCTCTGAAATGTACAGTTGGC 1199
Db 172567 TCCGCCACAAGATTTTATAGGATTTCTGCCATACACAGGCTCTGACA-----CCAGC 172619
Qy 1200 CTGAGACTATCTTCTTCTTCTCTAGGAGTGTGTCATGAACAATGCCACTGCACCTCGGG 1259
Db 172620 CTGAGTCTGTATCTTCTTCTTCTTGAGGAGTGTGTCATGAACAATGCCATCTGTACTCGGG 172679
Qy 1260 TCTATGAGAAGTGCAATGA 1279
Db 172680 TCTATGAGAAGGTACAATGA 172699

RESULT 5
AF181449 42874 bp DNA linear PRI 02-AUG-2002
LOCUS Homo sapiens chromosome 8 clone CTB-402C7 map 8q21-q23, complete
DEFINITION sequence.
ACCESSION AF181449
VERSION AF181449.3 GI:22091331
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 42874)
AUTHORS Blechschmidt,K., Schattevoy,R., Baumgart,C. and Rosenthal,A.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 42874)
AUTHORS Blechschmidt,K., Schilhabel,M., Schattevoy,R., Baumgart,C.,
Menzel,J., Weber,J., Korenberg,J.R. and Rosenthal,A.
Direct Submission
TITLE Submitted (23-AUG-1999) Genome Analysis, Institute of Molecular
JOURNAL Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 42874)
```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```
Schilhabel,M.B. and Platzter,M.
Direct Submission
Submitted (07-JUN-2001) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
4 (bases 1 to 42874)
Lagemann,D. and Platzter,M.
Direct Submission
Submitted (02-AUG-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
On Aug 2, 2002 this sequence version replaced gi:14327843.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H193
Center clone name: CTB-402C7
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42771 bases at least Q40
Consensus quality: 42810 bases at least Q30
Consensus quality: 42843 bases at least Q20
Quality coverage: 31.66x
```

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source

```
Location/Qualifiers
1..42874
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q21-q23, complete sequence."
/clone="CTB-402C7"
1..30
/note="single clone coverage , low quality region ,
CTB-402C7"
```

misc_feature

```
31..73
/note="low quality region , CTB-402C7"
```

misc_feature

```
7974..7983
/note="single stranded/single chemistry region"
```

misc_feature

```
14657..14770
/note="single stranded/single chemistry region"
```

misc_feature

```
14771..14851
/note="single stranded/single chemistry region"
```

misc_feature

```
14852..14981
/note="single stranded/single chemistry region"
```

misc_feature

```
14982..15030
/note="single stranded/single chemistry region"
```

misc_feature

```
27005..27140
/note="single stranded/single chemistry region"
```

misc_feature

```
27141..27748
/note="single stranded/single chemistry region"
```

misc_feature

```
42700
```

Db	6159	TTGCTTTGCTGCCACGCTCACAGTGAATCTTTGGCAAGCCACCACCACTGCACACATTTTTC	6211
Qy	983	ACATTTGATTAACTAAGGATATTTTCCAAAAACACATGAATATTTTAGAGATCATATCCA	1042
Db	6219	TATATCTTTTAAATAAGGAATATTTGCCATAAACCAAAAGATAAAGATATATTCAACACA	6278
Qy	1043	GTGCTTTAGTCTGCAGGGCAGCAATATATACATATAAACAACAGCAGCTCTAGGTCTT-	1101
Db	6279	TTAGTTTCTGTAAAAACATGTCATATTTATAAGCAAAAACAACAGCTTCTGCTCTCAAA	6338
Qy	1102	CTTGAGTTTGAATCCTGAGATGTTGTTTCTTGTAGTTGGTTTACAAGCGTTTATAGGA	1161
Db	6339	CCCGTGAATCTGAAAGAAATTTTCTCATCTATGAAGTAGATGATGATTTTCGTGGGA	6398
Qy	1162	TTCTGCC-----ACAACACATGCTCTGAATGTACAGTTGGCTCGGAGACTC	1208
Db	6399	CITTGATTTCTAATGTTTAAATACCACTGCTCTGGAATCTAAGGCTAACCTAACTCTT	6458
Qy	1209	T-----ATCTTTCTCTCTAGGAGTGTGTCATGAACAATGCCACTCGCACTCGGCTCTA	1263
Db	6459	TTAATATCTTTCTCTCTAGGAGTGTGTCATGAACAATGTCACTGCTACTCGGATCTA	6518
Qy	1264	TCAGAGGTGCAATGA	1279
Db	6519	TGAAAAAGTAGATAA	6534
RESULT 6			
AC022647	174316 bp DNA linear HTG 09-MAY-2001		
LOCUS	Homo sapiens clone RP11-28G10, WORKING DRAFT SEQUENCE, 4 unordered		
DEFINITION	pieces.		
ACCESSION	AC022647		
VERSION	AC022647.5 GI:13958526		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 174316)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.		
TITLE	Homo sapiens, clone RP11-28G10		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 174316)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Chospel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
COMMENT	On May 6, 2001 this sequence version replaced gi:13184149. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
----- Genome Center			
Center: Whitehead Institute/ MIT Center for Genome Research			
Center code: WIRB			
Web site: http://www-seq.wi.mit.edu			
Contact: sequence_submissions@genome.wi.mit.edu			

misc_feature	/note="low quality region , CTB-402C7"		
	42795..42874		
misc_feature	/note="single stranded/single chemistry region"		
	42846..42874		
unsure	/note="low quality region , CTB-402C7"		
	42850		
unsure	/note="CTB-402C7"		
	42855		
	/note="CTB-402C7"		
ORIGIN			
Query Match	23.1%; Score 296; DB 9; Length 42874;		
Best Local Similarity	59.3%; Pred. No. 5.1e-77; Indels 70; Gaps 10;		
Matches	721; Conservative 0; Mismatches 425;		
Qy	126	CCACAGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGAGCTGTATCA	185
Db	5327	CTACAGAGTGGGAATAGCTTTCGGAATAATGGCGCAATGGCCAAAGCCAGATGTATCA	5386
Qy	186	TTACGTGTGATGGCAACATCAGCGTCAAAACCGAGAGCAGATGAAGACAGCTGTGT	245
Db	5387	TCACITGTGATGGTAAAAAACCCTCACCATAAAACTGAGAGCACTTTGAAAAACAACACAGT	5446
Qy	246	TCCTTTTAACCTGGGAGAGAGTTTGATGAACGACAGCTGATGGCAGAAAACTGAGG	305
Db	5447	TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGG	5506
Qy	306	TCAGCTACACATCTGTGAAGCGACAGAGCTTCTAGATTTTACAGATTTAAATTTGCATTA	365
Db	5507	TCAGTGTGATGTTATGAATACACAGAGCTTCTAGATGATAGCTGTAT---CAAT	5563
Qy	366	ACATGCTGTACTACTGCCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTT	425
Db	5564	AACATTTTACTGTTTATAGGCAAGAACTTAATGAAGAGTTATTTTAT---GAATTTGAATTT	5622
Qy	426	TGATAAATTAGTAAAGTCCAGGACTAAGAAATGAACATCTTATGATGTTCTAGATC	485
Db	5623	TGTCAAATTAGCAAAAGTATCACTTCATCATAGAAATGGCATCTTTTATTAGCTACTAG	5682
Qy	486	GAAGACCATATGTTTATTTGTAACAAATCAGTATGATGGGTGGAGT---TCAGAG	541
Db	5683	GTTGAAACCAACAACTATTGTGTAATAAATCAATATGGTTAATGAAGTAGACTCAGAA	5742
Qy	542	AGGAAAGCGAAGACTTTTGGAGT-----GGTGGGTCTCTGGGGTCTCTTCACTT	595
Db	5743	AGGAGAAGGTGAACAAATGTTGATTAAAGAGGTTATGAGTCATGGAAACTCTTGTAAATG	5802
Qy	596	TGGAAGATGATGAACTAACTACCTGTATTTTTCAGACGGTCTGCACCTTCCAGACGG	655
Db	5803	TACTTGAAGATTAAAAACGTTTACTTTGTTTTCAGAGCTGTCTGCAACTTTTACAGATGG	5862
Qy	656	TGCCCTGTCCAGCACAGCAATGGGACGGGAAGGAGACAGATACAGAAAACTGAA	715
Db	5863	TGCATTTGGTTTACGATCAGAGTGGGATGGGAAGGAAGCAACATAACAGAAAAATTGAA	5922
Qy	716	GGATGGGAAGATGATCGTGGTGAAGATCAAGACACTGGCACCATGCTGGGATTTGGGCTG	775
Db	5923	AGATGGGAATTAGTGGTGGTAAAGTGTCAAACTGCTGTCTCAGTCAGCTTCTTGTGTG	5982
Qy	776	CAGCCACATTTGTATAAACCACTTCGGGTCAATGTTCTTTTAAACAGAGAAGGAACCTTA	835
Db	5983	CATTCTATGTTTCACTAACTGTTTCTATATCATTTGATCATT-----AACAGAACTCAGTTTG	6038
Qy	836	GGAGGACAACTACTGAAATACAGTTTGAAGAACGAGAGTCTCTCATTTGCTGAGCGCCCT	895
Db	6039	GAAGAAAAAAGCAAAATAACAGTTTAAATACTAGAACACTAATTTATTAAGAAAGTCTC	6098
Qy	896	TGTGGGACGGAAGTGTATGGG-----ATCCAGGATGTGGCTGCGACGAGCCTG----	947
Db	6099	AGTGGAGATAGAGAGTGAATGACATGACATAGAGAGGATTTGGCTGGAGTTGGGGGAGTTC	6158
Qy	948	-----AGAGCTGGCAGGCCACCGAGACGCCCTCTCTCTGT	982

----- Project Information

Center project name: L4751
Center clone name: 28_G.10

----- Summary Statistics

Sequencing vector: M13; M77815; 3% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 172808 bases at least Q40
Consensus quality: 173594 bases at least Q30
Consensus quality: 173848 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 174016; sum-of-contigs
Quality coverage: 8.3 in Q20 bases; agarose-fp
Quality coverage: 8.1 in Q20 bases.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 1470: contig of 1470 bp in length
* 1471 1570: gap of 100 bp
* 1571 2692: contig of 1122 bp in length
* 2693 2792: gap of 100 bp
* 2793 34878: contig of 32086 bp in length
* 34879 174316: gap of 100 bp
* 34979 174316: contig of 139338 bp in length.

FEATURES

source

1..174316
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-28G10"
/clone_lib="RPC1-11 Human Male BAC"

misc_feature

1..1470
/note="assembly_fragment"
clone_end:SP6
vector_side:left

misc_feature

1571..2692
/note="assembly_fragment"

misc_feature

2793..34878
/note="assembly_fragment"

misc_feature

34979..174316
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN

Query Match 23.1%; Score 296; DB 2; Length 174316;
Best Local Similarity 59.3%; Pred.No. 6.2e-77;
Matches 721; Conservative 0; Mismatches 425; Indels 70; Gaps 10;

QY 126 CCACAGGACTAGACTGGCTCTTAGAAGATGGCTGCCATGCCAAGCCAGACTGTATCA 185
Db 32378 CTACAGGATGGGAATAGCTTTGGGAAAATGGGGCAATGGCCAGCCAGATTGATCA 32437
QY 186 TTACGTGTGATGGCAACCAATCAGCGTCAAAACCCAGAGACAGCTGAAGACACTGTGT 245
Db 32438 TCACCTTGATGGTAAAAACCTCACCATAAAACCTGAGAGCACTTTGAAAACCAACACAGT 32497
QY 246 TCTCTTGTAACCTGGGAGAGAAGTTTGATGAACGACAGCTGATGGCAGAAAACTGAGG 305
Db 32498 TTCTTGTAACCTGGGAGAGAAGTTTGAAAACCCACAGCTGATGGCAGAAAACTCAGG 32557
QY 306 TCAGCTTACAACTACTGTGAAGCGACAGAGCTTCTAGATTTACAGATTTAAATTCATTA 365
Db 32558 TCAGTCGTGACATGTTATGAATTCACAGAAAGCTTCTAGATGATAGGCTGTAT---CAAT 32614
QY 366 ACAATGTCGTGTTACTTCTGCCAAGGCTGACTGAAAAAACTACTTTATGAGCTTGACTTT 425

Db 32615 AACATTTTACTGTTTATAGGCAAGAACTTAATGAAAAAGTTATTTTAT-GAATGAATTT 32673
QY 426 TGATAAATTTAGTAAAGTCCAGGACTAAGAAATGAAGACATCTTTATCAGTTTCTAGATC 485
Db 32674 TGTCAATTTAGCAAAAGTATCAACTTTCATCATAGATTTGGCATCTTTTATTAGCTACTAG 32733
QY 486 GAAAGACACATAGTTTGTATTGTGAACAAAATCAGTATGATGGGTGGAGT----TCAGAG 541
Db 32734 GTTGAACACCAACAACTATTGTGAATAAAATCAATATGGTTAATGAAGTAGACTCAGAA 32793
QY 542 AGGGAAGCGCAAGACTTTGTTGGAGT-----GGTGGGTCTGGGGGTTCTTTCACATT 595
Db 32794 AGGAGAGGTGAACCAAAATGTTGATTAAGGAGGTTATGAGTCATGGAACCTCTTGAATG 32853
QY 596 TGGAGATGATGAACCTAACCCTGTTATTTTTCAGAGCGTCTGCACCTTCCCAAGACGG 655
Db 32854 TACTTGGGAAGATTAAACGTTTACTTTTGTGTCAGACTGTCCTCAACTTTTACAGATGG 32913
QY 656 TGCCCTGCTCAGCACCAACCAATGGGACGGGAAGAGACGATAAACAAGAAAACTGAA 715
Db 32914 TGCATTGCTCAGCATCAGGATGGGATGGGAAGGAAGCAATTAACAAGAAAAATTGAA 32973
QY 716 GGATGGGAAGATGATCGTGGTGAGCATCAAAAGCACTGGCACCATGCTGGGATTTGGCCCTG 775
Db 32974 AGATGGGAATTTAGTGGTGGTAAGTGTCAAACTGCTGTCTCAGTCAGCTTCTTGTGTG 33033
QY 776 CAGCCACAGTTGTCTATAACCACTTCGGGTCTATTGGTCTTTTAAACAAGAGAAGAACTTA 835
Db 33034 CATTCATAGTTTACATAAATCTGTTCTATATCATTTGATCAATT-----AACAGAATCAGTTG 33089
QY 836 GGAGGACAATATCTGAAAAATAACAAGTTAGAAAACGAGAGTCTCTCATTTGCTAGGCGACCT 895
Db 33090 GAAGAAAAAAGCAAAATAACAAGTTAAATACTAGAACACTAAATATTATTAAGAAAGTCT 33149
QY 896 TGTGGGACGGGAAGATGATGGG-----ATCCAGATGTGCTGTCAGCAGACAGCCTG---- 947
Db 33150 AGTGGAGATAGAGAAGTCAATGATAGGAGAGGATTTGGCTGGAGTTGGGGGGGAGTTC 33209
QY 948 -----AGAGCTGGCAGGCCACCGAGCAGCCCTCTCTCTGT 982
Db 33210 TTGCTTTGCTGCCAGCTCAGAGTGAATCTTGGGGAAGCCACAACTGCACACATTTTC 33269
QY 983 ACATTGATTTAAGTAAGGATATTTGCCAAACACATGAATAAATTAGAGATCATATCA 1042
Db 33270 TATATCTTTAAAAATAAGGATATTTGCCATAACCAAAAGATAAAGATATATTCAACACA 33329
QY 1043 GTGCTTTAGTCTGCGAGGCGCAATATACATATAAACAACACAGCAGCTTAGTCTTT- 1101
Db 33330 TTAGTTTTCTGTAAACATGTCATATTATAGCAAAACACAGCTTCTGGCTTCTCTCAAA 33389
QY 1102 CTTGAGTTTGAATCCTGAGATGTGTTTTTCTTTAGTTAGTTTGAAGCGTTTATAGGA 1161
Db 33390 CCGTGAATTTGAAAGAAATTTTCTCATCTATGAGTAGATTAAGTATTTCTGGGA 33449
QY 1162 TTCTGCC-----ACAACACATGCTCTGAAATGTPACAGTTGGCTGAGACTC 1208
Db 33450 CTTTGAATCTTAATGTTTTTAATACCACCTGCTCTGGAATCTAAGCTAAACCTAATCTTT 33509
QY 1209 T-----ATCTTTTCTCTAGGAGTGTCTATGAACAATGCCACCTGCACCTCGGGTCTA 1263
Db 33510 TTAATATCTTTCTTTCTTAGGAGTGTCTATGAACAATGTCACCTGTACTCGGATCTA 33569
QY 1264 TGAGAAGGTGCAATCA 1279
Db 33570 TGAAAAAGTAGAATAA 33585

RESULT 7

AC009902/c

LOCUS

DEFINITION

AC009902

ACCESSION

VERSION

217249 bp DNA linear PRI 18-OCT-2002
Homo sapiens chromosome 8, clone RP11-361E6, complete sequence.

AC009902.13 GI:24111046

KEYWORDS	HTG.	Direct Submission	TITLE
SOURCE	Homo sapiens (human)	Submitted (18-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	JOURNAL
ORGANISM	Homo sapiens	On Oct 18, 2002 this sequence version replaced gi:18250063.	COMMENT
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html	
AUTHORS	1 (bases 1 to 217249)		
TITLE	Birken, B., Nusbaum, C. and Lander, E.		
REFERENCE	Homo sapiens chromosome 8, clone RP11-363E6		
AUTHORS	2 (bases 1 to 217249)		
	Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehocsky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 217249)		
AUTHORS	Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Govette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihoval, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (05-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	4 (bases 1 to 217249)		
AUTHORS	Birken, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihoval, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	On Oct 18, 2002 this sequence version replaced gi:18250063.		
AUTHORS	All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
	Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence.submissions@genome.wi.mit.edu ----- Project Information Center project name: L2420 Center clone name: 363_E_6 ----- Location/Qualifiers 1. 217249 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="8" /map="8" /clone="RP11-363E6" /clone_lib="RP11-363E6" /complement(335..434) /rpt_family="MIR" 1670..2741 /rpt_family="LIMA8" 3191..3298 /rpt_family="GA-rich" 5349..5647 /rpt_family="AluY" 5698..6174 /rpt_family="MLT1J" 7870..7929 /rpt_family="LIMB6" 7932..8370 /rpt_family="MER65C" 8377..9253 /rpt_family="LIMB6" 9255..9278 /rpt_family="AT-rich" 9281..9566 /rpt_family="AluJo" complement(10067..10362) /rpt_family="AluSx" 10883..11054 /rpt_family="MIR" complement(11067..11376) /rpt_family="AluSx" complement(11691..11917) /rpt_family="AluSg/x" 12128..12175 /rpt_family="AT-rich" 13088..13114 /rpt_family="(CAAA)n" 13768..13935 /rpt_family="MIR" complement(15232..15537) /rpt_family="AluSx" 17458..17504 /rpt_family="CT-rich" 17689..17882 /rpt_family="MIR" 18317..18761 /rpt_family="MLT1C" 19656..19752 /rpt_family="CT-rich" 20146..20278 /rpt_family="FLAM_C" 20313..20534 /rpt_family="AluJb"		

repeat_region	complement(20860. .20968) /rpt_family="LIME3A"
repeat_region	20978. .21273 /rpt_family="AluJb"
repeat_region	complement(21291. .21755) /rpt_family="LIME3"
repeat_region	complement(21768. .21875) /rpt_family="LIME3"
repeat_region	complement(21941. .22386) /rpt_family="MER97C"
repeat_region	22655. .22740 /rpt_family="L2"
repeat_region	22857. .28882 /rpt_family="L2"
repeat_region	28905. .30215 /rpt_family="LIP3"
repeat_region	28905. .30215 /rpt_family="LIPB3"
repeat_region	complement(30546. .30638) /rpt_family="L2"
repeat_region	complement(30815. .30948) /rpt_family="MLT1H"
repeat_region	31247. .31553 /rpt_family="AluSx"
repeat_region	31824. .32440 /rpt_family="MER67C"
repeat_region	complement(32898. .33003) /rpt_family="MIR"
repeat_region	33235. .33450 /rpt_family="MIR"
repeat_region	complement(33480. .33730) /rpt_family="L3"
repeat_region	complement(33731. .34102) /rpt_family="MSTB"
repeat_region	complement(34107. .34423) /rpt_family="L3"
repeat_region	complement(34424. .34738) /rpt_family="AluSx"
repeat_region	complement(34739. .34873) /rpt_family="L3"
repeat_region	complement(34874. .35339) /rpt_family="MLT1J"
repeat_region	complement(35340. .35425) /rpt_family="L3"
repeat_region	complement(35871. .36022) /rpt_family="MER5A"
repeat_region	36047. .36152 /rpt_family="L2"
repeat_region	complement(36244. .36464) /rpt_family="MIR"
repeat_region	37174. .37283 /rpt_family="MIR"
Query Match 23.1%; Score 296; DB 9; Length 217249; Best Local Similarity 59.3%; Pred. No. 6.4e-77; Matches 721; Conservative 0; Mismatches 425; Indels 70; Gaps 10;	
QY	126 CCACAGAGTAGACTGGCTCTTAGAAGATGCTGCCATGCCAAGCCAGACTGATCA 185
Db	71377 CTACAGAGTGGAAATAGCTTTGCGAATAATGGCGCAATGGCAAGCCAGATTGATCA 71318
QY	186 TTACGTGTGATGCCAACACATCACGGTCAAAACCGAGAGACAGTGAAGACACTGTGT 245
Db	71317 TCACCTGTGTGTAATAAACCTCACCATATAAAACTGAGAGCACTTTTGAAACCAACACACT 71258
QY	246 TCTCTTGTAACTGGGAGAGAGTTTGATGAACGACGACTGATGGCAGAAAACTGAGG 305
Db	71257 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACGACTGATGGCAGAAAACTCAGG 71198
QY	306 TCAGCTACACACTACTGTGAAGCGACAGAGCTTCTAGATTTCACAGATTAAATTGCATTA 365
Db	71197 TCAGTCGTGACATCTTATGAATACACAGAGCTTCTAGAATGATAGCTGTAT--CAAT 71141
QY	366 ACAATCTCTGTACTTACTGCCAAGGGCTGACTGAAAAAACTACTTTATCGAGTTGACTTT 425

Db	71140 AACATTTTACTGTTTTATAGGCAAGAACTTAATGAAAAAGTTATTTTAT-GAATTGAATTT 71082
QY	426 TGATAAATTTAGTAAAGTCCCAGGACTAAGAAATGAAGACACTCTTTATGAGTTTCTAGATC 485
Db	71081 TGTCAAATTAGCAAAAGTATCAACTTCATCATGAATTTGGCATCTTTTATTAGCTACTAG 71022
QY	486 GAAAAGCACATAGTTTGTATGTGAACAAAAATCAGTATGATGGGTGGAGT----TCAGAG 541
Db	71021 GTTGAANAACCAAACTATTGTGAATAAAATCAATATGGTTTAATGAAGTAGACTCAGAA 70962
QY	542 AGGGAAGCGGAGAGACTTTGTTGGAGT-----GGTGTGGTCTCTGGGGTTCCTTCACATT 595
Db	70961 AGGAGAGGTGAAAAAAAATGTTGATTAAGGAGGTTTATGAGTCATGGAACCTCTTGAATG 70902
QY	596 TGGAGATGATGAACCTAACTACCTCCCTGTATTTTTCAGACGGTCTGCACCTCTCCAAGACGG 655
Db	70901 TACTTGGAGATTAAACGTTTACTTTTGTTTTTCAGACTGTCTGCAACTTTTACAGATGG 70842
QY	656 TGCCCTGGTCCAGCACCCAGCAATGGGCGGGAAGAGAGACGATPAAACAAGAAAACTGAA 715
Db	70841 TGCATTGGTTCAGCATCAGGATGGGATGGGAAGGAAAGCAACAATAACAAGAAAAATTGAA 70782
QY	716 GGAATGGGAAGATGATCGTGTGAGCATCAAGACACTGGCACCCTCTGGGATTTGGCCCTG 775
Db	70781 AGATGGGAAATTTAGTGGTGGTAAAGTGTCAAACTGCTGTCTCAGTCAGCTTCTTTGTGTG 70722
QY	776 CAGCCACAGTTGTCTATAACCACTTCGGGTCTATTGGTCTCTTAAACAAGAGAAGAACTTA 835
Db	70721 CATTCATAGTTTACATATACTGTTCTATATCATTTGATCAIT-----AACAGACTCAGTTTG 70666
QY	836 GGAGGACAAATCTGAAAAATAACAAGTTAGAAACGAGAGTCTCTCATTTGCTAGGCGACCCCT 895
Db	70665 GAAGAAAAAAGCAAAATAACAAGTTAAATACTAGAACACTAAATTAATTAAGAAAACTCT 70606
QY	896 TGTGGGACGGAGAAAGTGAATGG-----ATCCAGAGATGTGGCTGCAGACAGACCTG---- 947
Db	70605 AGTGAGATAGAGAAAGTGCATGACATAGGAGAGGATTTGGCTGGAGTTGGGGGGAGTTTC 70546
QY	948 -----AGAGCTGGCAGGCGCCAGCAGCAGCCCTCTCTCTGT 982
Db	70545 TTGCTTTTGTGCCACGTCACAGTGAATCTTGGCAAGCCACCAACTGCACACATTTTC 70486
QY	983 ACAITGATTTAAGTAAGGATAATTTGCCAAAAACACATGAATAATTTAGAGATCATATCCA 1042
Db	70485 TATACTCTTAAATAAAGGAATATTTGCCATAACCAAAAGATAAAGATATATTCAACACA 70426
QY	1043 GTGCTTTAGTCTGAGGGCAGCAATATACATATAAACAACACAGCAGCTCTAGTCTTT- 1101
Db	70425 TTAGTTTTCTGTAAAAACATGCATATTATAAGCAAAACACAGCTTCTGGCTTCTCTCAAA 70366
QY	1102 CTTGAGTTTGAATCCTGAGATGTGTTTTCTTTAGTTGGTTGTTTACAAGCGTTTATAGGA 1161
Db	70365 CCGGTGAATCTGAAAGAAATTTTCTCCATCTATGAGTAGATTAGTGATTTCTGGGA 70306
QY	1162 TTCTGCC-----ACAAACATGCTCTGAAATGTPACAGTTGGCTGTGAGCTC 1208
Db	70305 CTTTGATCTTAATGTTTTTAATACCACCACTGCTCTGGAATCTAAGGCTAAACCTAACTCTT 70246
QY	1209 T-----ATCTTTCTCTCTAGGAGTGTGTCAATGAACATGCCACTGCACCTCGGGCTTA 1263
Db	70245 TTAATATCTTCTCTCTTAGGAGTGTGTCTATGAACAATGTCACCTGTACTCGGATCTA 70186
QY	1264 TGAAGAGGTGCAATGA 1279
Db	70185 TGAANAAGTAGAATAA 70170

RESULT 8
CO777551
LOCUS
DEFINITION Sequence 1237 from Patent EP1394274.
ACCESSION CO777551
VERSION CO777551.1 GI:45380535

933 bp DNA linear PAT 11-MAR-2004

TITLE
 JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 JOURNAL

Schnersch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 713)
 Strausberg, R.
 Direct Submission
 Submitted (31-JAN-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clon distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 8 Row: m Column: 7
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 6754449.
 Location/Qualifiers
 1..713
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="mix FVB/N, C57BL/6J"
 /db_xref="taxon:10090"
 /clone="MGC:5786 IMAGE:3490535"
 /tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
 old, gross tissue."
 /clone_lib="NCI CGAP_Mam5"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..713
 /gene="Fabp5"
 /note="synonym: mall"
 /db_xref="LocustID:16592"
 /db_xref="MGI:101790"
 36..443
 /gene="Fabp5"
 /codon_start=1
 /product="fatty acid binding protein 5, epidermal"
 /protein_id="AAH02008.1"
 /db_xref="GI:12805103"
 /db_xref="LocustID:16592"
 /db_xref="MGI:101790"
 /translation="MASLKDLGKRLMESHGFSEYMKELGVGLAIRYMAAWAKPDCI
 ITCDDNNITVETSTKTVFSCNLGKPFDETTADGRKTETVCTFQDGLVQHQWDG
 KESTITRLKDKGMIVCEVMNNATCTRYEVQV"

ORIGIN
 Query Match 13.7%; Score 175; DB 10; Length 713;
 Best Local Similarity 94.8%; Pred. No. 7.8e-41;
 Matches 181; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
 |||
 Db 113 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 172
 |||
 QY 190 GTGTGATGGCAACAACATCAGCGTCAAAACCGGAGACACAGTGAGACGACTGTGTTCTC 249
 |||
 Db 173 GTGTGATGGCAACAACATCAGCGTCAAAACCGGAGACACAGTGAGACGACTGTGTTCTC 232
 |||
 QY 250 TTGTACCTGGGAGAGAAAGTTTGTATCAAAACGACAGCTGATGGCAGAAAAAAGCTGAGTCTAG 309
 |||

Db 233 TTGTAACTGGGAGAGAGTTTGTATGAACACAGCTGATGGCAGAAAAAAGCTGAGACGCT 292
 |||
 QY 310 CTACAAATAC 320
 |||
 Db 293 CTGCACCTTCC 303
 |||

RESULT 11
 AC125321 196150 bp DNA linear ROD 25-NOV-2003
 LOCUS Mus musculus BAC clone RP23-371B13 from chromosome 9, complete
 DEFINITION
 AC125321
 VERSION AC125321.4 GI:37951467
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 196150)
 Berghoff, A., Haglund, K., Kozlowicz, A., Bielicki, L. and Meyer, R.
 The sequence of Mus musculus BAC clone RP23-371B13
 Unpublished (2001)
 REFERENCE
 2 (bases 1 to 196150)
 Wilson, R.
 Sequencing of Mus musculus
 Unpublished (2001)
 3 (bases 1 to 196150)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (22-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 196150)
 Wilson, R.K.
 Direct Submission
 Submitted (06-AUG-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 196150)
 Wilson, R.K.
 Direct Submission
 Submitted (24-OCT-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 6 (bases 1 to 196150)
 Wilson, R.
 Direct Submission
 Submitted (25-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Oct 24, 2003 this sequence version replaced gi:33457280.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@watson.wustl.edu
 ----- Summary Statistics

 Center project name: M_BA0371B13

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,

QY 190 GTGTGATGCGCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTTCTC 249
 |||||
 Db 182112 GTGTGATGCGCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTTCTC 182171
 |||||

QY 250 TTGTAACTCTGGGAGAGAGTTTGTGTAACGACGACGCTGATGCGCAAGAACTGAG 304
 |||||
 Db 182172 TTGTAACTCTGGGAGAGAGTTTGTGTAACGACGACGCTGATGCGCAAGAACTGAG 182226
 |||||

RESULT 12
 AC124432/c
 LOCUS AC124432 160728 bp DNA linear ROD 27-NOV-2003
 DEFINITION Mus musculus BAC clone RP24-243L15 from chromosome 18, complete
 sequence.
 AC124432
 VERSION AC124432.5 GI:31194018
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 160728)
 AUTHORS Swearingen-Shahid, S. and Meyer, R.
 TITLE The sequence of Mus musculus BAC clone RP24-243L15
 JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 160728)
 AUTHORS Wilson, R.
 TITLE Sequencing of Mus musculus
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 160728)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 160728)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (02-APR-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 160728)
 AUTHORS Wilson, R.K.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-2003) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 160728)
 AUTHORS Wilson, R.
 TITLE Direct Submission
 JOURNAL Submitted (27-NOV-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On May 30, 2003 this sequence version replaced gi:29469630.

 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu
 Contact: submissions@watson.wustl.edu

 Summary Statistics

 Center project name: M_BB0243L15

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
 http://genome.wustl.edu

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and
 coworkers (http://www.chori.org) from male C57BL/6J mouse spleen
 and/or brain genomic DNA. The clone and detailed information can be
 obtained from Pieter de Jong and coworkers at http://www.chori.org

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES	source	Location/Qualifiers
	1..160728	/organism="Mus musculus"
		/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		/chromosome="18"
		/map="18"
		/clone="RP24-243L15"
		/clone_lib="RPCI-24"
repeat_region	1629..1683	/rpt_family="L1"
repeat_region	3465..3530	/rpt_family="MIR"
repeat_region	3709..3767	/rpt_family="MER1_type"
repeat_region	3768..4484	/rpt_family="RMR6A"
repeat_region	4493..4545	/rpt_family="B4"
repeat_region	6368..6897	/rpt_family="ERV1"
repeat_region	7286..7388	/rpt_family="MaLR"
repeat_region	7625..7848	/rpt_family="RMR6A"
repeat_region	7978..8124	/rpt_family="Alu"
repeat_region	8127..8519	/rpt_family="RMR6A"
repeat_region	8525..8668	/rpt_family="Alu"
repeat_region	8742..9568	/rpt_family="RMR19A"
repeat_region	9569..9649	/rpt_family="RMR6A"
repeat_region	9655..10034	/rpt_family="L1"
repeat_region	10065..10342	/rpt_family="ERVK"
repeat_region	10585..10634	/rpt_family="RMR6A"
repeat_region	10663..10852	/rpt_family="RAL_RN"
repeat_region	10695..11211	/rpt_family="ERV1"
repeat_region	11212..11264	/rpt_family="ERV1"
repeat_region	11284..11521	/rpt_family="ERV1"
repeat_region	12965..13010	/rpt_family="B4"
repeat_region	13012..13075	/rpt_family="ERV1"
repeat_region	13348..13618	/rpt_family="ERV1"
repeat_region	14328..14663	/rpt_family="ERV1"
repeat_region	14665..14713	/rpt_family="tRNA-Arg-AGG"
repeat_region	14716..14867	

```
repeat_region /rpt_family="ERV1"
14905..15461
/rpt_family="ERV1"
15462..15953
/rpt_family="RMR6A"
16095..16172
/rpt_family="Alu"
16295..16477
/rpt_family="B2"
16743..16827
/rpt_family="ERV1"
16836..17256
/rpt_family="ERVK"
18595..18691
/rpt_family="B4"
18776..18922
/rpt_family="Alu"
20784..20974
/rpt_family="B2"
21648..21835
/rpt_family="B2"
21846..21973
/rpt_family="Alu"
23208..23348
/rpt_family="Alu"
24837..24910
/note="Sequence derived from PCR product of project DNA."
repeat_region 25025..25233
/rpt_family="L1"
25075..25085
/note="Sequence derived from one plasmid subclone."
repeat_region 25469..25619
/rpt_family="Alu"
25622..25734
/rpt_family="L1"
25769..25929
/rpt_family="B4"
26971..27138
/rpt_family="ERV1"
27706..27873
/rpt_family="B2"
27874..28022
/rpt_family="Alu"
28023..28066
/rpt_family="B2"
28362..28381
/rpt_family="B2"
28382..28531
/rpt_family="Alu"
28532..28700
/rpt_family="B2"
29417..29766
/rpt_family="MaLR"
31583..31941
/rpt_family="MaLR"
32560..32657
/rpt_family="MaLR"
32658..32798
/rpt_family="Alu"
32944..33341
/rpt_family="ERVK"
34057..34154
/rpt_family="Alu"
34077..34168
/rpt_family="B4"

Query Match 12.9%; Score 165.4; DB 10; Length 160728;
Best Local Similarity 91.6%; Pred. No. 1.3e-37;
Matches 175; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCATGGCCAGCCAGACTGTATCATTTAC 189
Db 22728 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCATGGCCAGCCAGACTGTATCATTTAC 22669
```

```
QY 190 GTGTGATGGCAACAACATCACCGGTCAAAACCGAGAGACACAGTGAAGACGACTGTGTCTC 249
Db 22668 GTGTGATGGCAACAACATCACCGGTCAAAACCGAGAGACACAGTGAAGACACCGGTGTCTC 22609
QY 250 TTGTAACTGGGAGAGAACTTTGATGAAACGACAGCTGATGGCAGAAAACCTCAGGTTCAG 309
Db 22608 TTGTAACTGGGAGAGAACTTTGATGAAACGACAGCTGATGGCAGAAAACCGAGAGGT 22549
QY 310 CTACACATAC 320
Db 22548 CTGACCTTCC 22538

RESULT 13
AC147992 239075 bp DNA linear ROD 10-JUL-2004
LOCUS Mus musculus BAC clone RP24-384K8 from chromosome 18, complete
DEFINITION sequence.
ACCESSION AC147992
VERSION AC147992.3 GI:49170304
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Tomlinson, C., Bielicki, L. and Meyer, R.
The sequence of Mus musculus BAC clone RP24-384K8
Unpublished (2001)
REFERENCE 3 (bases 1 to 239075)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 239075)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (24-JUN-2004) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 5 (bases 1 to 239075)
AUTHORS Wilson, R.K.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2004) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 24, 2004 this sequence version replaced gi:47131351.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Summary Statistics
-----
Center project name: M_BB0384K08
-----
```

NOTICE:

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren,

Matches 175; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189

Db 218789 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 218848

QY 190 GTGTGATGGCAACAACATCATCGGTCAAAACCGAGGACGACAGTGAAGACGACTGTGTTCTC 249

Db 218849 GTGTGATGGCAACAACATCATCGGTCAAAACCGAGGACGACAGTGAAGACGACTGTGTTCTC 218908

QY 250 TTGTAACTCTGGGAGAGAGTTTGTATGAACACGACAGCTGATGGCAGAGAAAACCTGAGTTCAG 309

Db 218909 TTGTAACTCTGGGAGAGAGCTTTGATGAACACGACAGCTGATGGCAGAGAAAACCTGAGTTCAG 218968

QY 310 CTACAACATAC 320

Db 218969 CTGCACCTCC 218979

RESULT 14

AL954662/c

LOCUS

DEFINITION Mouse DNA sequence from clone RP23-348L16 on chromosome 2, complete sequence.

ACCESSION AL954662

VERSION AL954662.13 GI:37718670

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Brown, J.

AUTHORS Direct Submission

TITLE Submitted (17-OCT-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

JOURNAL humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk

COMMENT On Oct 17, 2003 this sequence version replaced gi:33666998. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun align has only been used where it has a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep

from the RFLP-23 Mouse BAC Library constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

FEATURES

source Location/Qualifiers

1. .132427

ORIGIN

Query Match 12.9%; Score 165; DB 10; Length 132427;

Best Local Similarity 92.1%; Pred. No. 1.7e-37;

Matches 174; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189

Db 14768 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 14709

QY 190 GTGTGATGGCAACAACATCATCGGTCAAAACCGAGGACGACAGTGAAGACGACTGTGTTCTC 249

Db 14708 GTGTGATGGCAACAACATCATCGGTCAAAACCGAGGACGACAGTGAAGACGACTGTGTTCTC 14649

QY 250 TTGTAACTCTGGGAGAGAGTTTGTATGAACACGACAGCTGATGGCAGAGAAAACCTGAGTTCAG 309

Db 14648 TTGTAACTCTGGGAGAGAGTTTGTATGAACACGACAGCTGATGGCAGAGAAAACCTGAGTTCAG 14589

QY 310 CTACAACAT 318

Db 14588 CTGCACCTT 14580

RESULT 15

AL939157

LOCUS

DEFINITION Mus musculus BAC clone RP24-152G20 from chromosome 18, complete sequence.

ACCESSION AC139157

VERSION AC139157.4 GI:40288242

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Swearingen-Shahid, S., Shahid, S., Kozlowicz, A., Bielicki, L. and Meyer, R.

AUTHORS The sequence of Mus musculus BAC clone RP24-152G20

TITLE Unpublished (2001)

JOURNAL Unpublished (2001)

REFERENCE 2 (bases 1 to 133869)

AUTHORS Wilson, R.

TITLE Sequencing of Mus musculus

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 133869)

AUTHORS McPherson, J.D. and Waterston, R.H.

TITLE Direct Submission

JOURNAL Submitted (24-JAN-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 133869)

AUTHORS Wilson, R.K.

TITLE Direct Submission

JOURNAL Submitted (24-SEP-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 133869)

AUTHORS Wilson, R.K.

TITLE Direct Submission

JOURNAL Submitted (22-DEC-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 6 (bases 1 to 133869)

AUTHORS Wilson, R.

TITLE Direct Submission

JOURNAL Submitted (01-JAN-2004) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT On Dec 22, 2003 this sequence version replaced gi:35069405.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: submissions@wustl.edu
 ----- Summary Statistics

 Center project name: M_BB0152G20

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC130218.

FEATURES

```

source
1..133869
   /organism="Mus musculus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10090"
   /chromosome="18"
   /map="18"
   /clone="RP24-152G20"
   /clone_lib="RPCI-24"
10978..11100
   /note="Sequence derived from one plasmid subclone."
10988..11185
   /note="Unresolved simple sequence repeat."
24188..24349
   /note="Unresolved simple sequence repeat."
113859..114855
   /note="Unresolved tandem repeat."
113860..114647
   /note="CpG island (%GC=71.2, o/e=0.74, #CpGs=97)"
133392..133869
   /note="Unresolved tandem repeat."

```

ORIGIN

```

Query Match      12.7%; Score 162.2; DB 10; Length 133869;
Best Local Similarity 90.6%; Pred. No. 1.2e-36;
Matches 173; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 130 AGGAGTAGACTGGCTCTTAGGAGATGGCTGCCAGCCAGACGACTGATCATTTAC 189
      |||||
DB 86505 AGGAGTAGACTGGCTCTTAGGAGATGGCTGCCAGCCAGACGACTGATCATTTAT 86564
      |||||
QY 190 GTGTGATGGCAACACATCACCGTCAAAACCCAGAGCACAGTGAAGACGACTGTGTTC 249
      |||||
DB 86565 TTGTGATGGCAACACATCACCTTCAAAACCCAGAGCACAGTGAAGACGCGGTGTTCTC 86624
      |||||
QY 250 TTGTAACCTGGAGAGAGAGTTTGTGATAAAGCAGCAGCTGATGGCAGAAAAAAGTGGTCTAG 309
      |||||
DB 86625 TTGTACCTCTGGAGAGAGAGTTTGTGATAAAGTACGGCTGATGGCAGAAAAAAGTGGACCGGT 86684
      |||||

```

QY 310 CTACAAACATAC 320
 |||||
 DB 86685 CTGCACCTTCC 86695

Search completed: July 12, 2005, 17:41:12
 Job time : 5913.77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 18:01:50 ; Search time 5881.03 Seconds
(without alignments)
8878.154 Million cell updates/sec

Title: US-09-788-074-2

Perfect score: 1279

Sequence: 1 aatgggagcaacatgctagc.....tctatgagaagtgcattga 1279

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/ECTUS1_COMB.seq:*

2: /cgn2_6/ptodata/1/pna/ECTUS2_COMB.seq:*

3: /cgn2_6/ptodata/1/pna/ECTUS_COMB.seq:*

4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*

5: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*

6: /cgn2_6/ptodata/1/pna/US080_COMB.seq:*

7: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*

8: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*

9: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*

10: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*

11: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*

12: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*

13: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*

14: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*

15: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*

16: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*

17: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*

18: /cgn2_6/ptodata/1/pna/US092A_COMB.seq:*

19: /cgn2_6/ptodata/1/pna/US092B_COMB.seq:*

20: /cgn2_6/ptodata/1/pna/US093A_COMB.seq:*

21: /cgn2_6/ptodata/1/pna/US093B_COMB.seq:*

22: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*

23: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*

24: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*

25: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*

26: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*

27: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*

28: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*

30: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*

31: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*

32: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

33: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*

34: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*

35: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*

36: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*

37: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*

38: /cgn2_6/ptodata/1/pna/US098D_COMB.seq:*

39: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*

40: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*

41: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*

42: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*

43: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq:*

45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq:*

46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*

47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*

48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*

49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*

50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*

51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:*

53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:*

54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq:*

55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq:*

56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq:*

57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq:*

58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:*

59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:*

60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:*

61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:*

62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:*

63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq:*

64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq:*

65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:*

66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:*

67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:*

68: /cgn2_6/ptodata/1/pna/US110_COMB.seq:*

69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*

70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*

71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*

72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*

73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*

74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*

75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*

76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*

77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*

78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*

79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*

80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*

81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*

82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*

83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*

84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*

85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*

86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*

87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*

88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*

89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*

90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*

91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*

92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq:*

93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq:*

94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*

95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*

96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*

97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*

98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*

99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*

100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*

101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*

102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*

103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*

104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*

105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*

106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*

107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*

108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*

109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*

110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*

111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq:*

113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq:*

114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq:*

115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq:*

116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq:*


```
Db 721 GGAAGATGATCGTGGTGAGCATCAAGCACTGGCACCATGCTGGGATTGGGCGTCGAGCC 780
Qy 781 ACAGTTGTCTAATCAACCACTTGGGTCTATGGTCTTTTAAACAAGAGAAAGAACTTAGGAGG 840
Db 781 ACAGTTGTCTAATCAACCACTTGGGTCTATGGTCTTTTAAACAAGAGAAAGAACTTAGGAGG 840
Qy 841 ACAATACCTGAAATACCAAGTTAGAAACGAGAGTCTCTATGCTGAGGCGAGCCCTTGTGG 900
Db 841 ACAATACCTGAAATACCAAGTTAGAAACGAGAGTCTCTATGCTGAGGCGAGCCCTTGTGG 900
Qy 901 GGACGGAGAGTGTGGGATCCAGGATGCTGGCTGAGCAGAGCCCTGAGAGCTGGCAGGC 960
Db 901 GGACGGAGAGTGTGGGATCCAGGATGCTGGCTGAGCAGAGCCCTGAGAGCTGGCAGGC 960
Qy 961 CACCGAGCAGCCCTCTCTGGTACATTTGATTTAAGTAAGGATATTTGCGCAAAACACATG 1020
Db 961 CACCGAGCAGCCCTCTCTGGTACATTTGATTTAAGTAAGGATATTTGCGCAAAACACATG 1020
Qy 1021 AATAATTTAGAGATCATATCCAGTGTCTTGAATCTGAGATGTGGTTTTCTGTTAGGT 1080
Db 1021 AATAATTTAGAGATCATATCCAGTGTCTTGAATCTGAGATGTGGTTTTCTGTTAGGT 1080
Qy 1081 AAAACAGCAGCTCTAGTCTCTTGAATCTGAGATGTGGTTTTCTGTTAGGT 1140
Db 1081 AAAACAGCAGCTCTAGTCTCTTGAATCTGAGATGTGGTTTTCTGTTAGGT 1140
Qy 1141 TGGTTACAAAGCGTTTATAGGATTTGCCCAACAACATGCTCTGAAATGACAGTTGGCC 1200
Db 1141 TGGTTACAAAGCGTTTATAGGATTTGCCCAACAACATGCTCTGAAATGACAGTTGGCC 1200
Qy 1201 TGAGACTCTATCTTCTCTAGGAGTGTGTCATGAACATGCCACCTGCACTGGGT 1260
Db 1201 TGAGACTCTATCTTCTCTAGGAGTGTGTCATGAACATGCCACCTGCACTGGGT 1260
Qy 1261 CTATGAGAGGTGCAATGA 1279
Db 1261 CTATGAGAGGTGCAATGA 1279

RESULT 2
US-60-466-412-85481
; Sequence 85481, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85481
; LENGTH: 16239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-466-412-85481

Query Match 23.2%; Score 296.8; DB 116; Length 16239;
Best Local Similarity 59.2%; Pred. No. 7.4e-76;
Matches 720; Conservative 2; Mismatches 424; Indels 70; Gaps 10;

Qy 126 CCACAGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCA 185
Db 8820 CTACAGGAGTGGAAATAGCTTTGCGAATAATGGCGCAATGGCCAGCCAGATGTATCA 8879
Qy 186 TTACGTGTGATGGCAACAACATCAGGTCAAAACCGAGAGACAGTGAAGACGACTGTGT 245
Db 8880 TCACCTGTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCAGCTTTGAAAAACAACAGT 8939
Qy 246 TCTCTTGTAACTGGGAGAGAGTTTGAATGAACGACAGCTGATGCGAGAAAACTGAGG 305
Db 8940 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGG 8999
```

```
Qy 306 TCAGCTACACATACATCTGTGAAGCGACAGAAAGCTTCTAGATTTACAGATTAATAATTGCATTA 365
Db 9000 TCAGTCGTGACATGTTATGAATAACAGAAAGCTTCTAGAAATGATAGGCTGTAT---CAAT 9056
Qy 366 ACAATGTCTGTACTTACTGCGCAAGGCTGACTGAAAAAACTACTTTATCGAGTTGACTTTT 425
Db 9057 AACATTTTACTCTTTTATAGGCAAGAACTTAATGAAAAAGTTATTTTAT---GAATGAAATTT 9115
Qy 426 TGATAAATTTAGTAAAGTCCCAAGGACTAAGAAATGAAGACATCTTATGAGTTTCTAGATC 485
Db 9116 TGTCAAATTTAGCAAAAGTATCAACTTCATCATAGAAATGGCATCTTTTATTAGTACTAG 9175
Qy 486 GAAAAGCAGATGTTGTTATTTGAAACAAAATCAGTATGATGGGTGGAGT----TCAGAG 541
Db 9176 GTTGAAGAACCAAAACTATTTGTAATAAAATCAATAATGSGTTAATGAAGTAGACTCAGAA 9235
Qy 542 AGGGAAGGCGAAGACTTGTTCGAGT-----GGTGTGGTCTCGGGGTTCCTTCACATT 595
Db 9236 AGGAGAGGTGAACAAAATGTTGATTTAAGGAGGTTATGAGTCAATGAAAATCTCTGTAATG 9295
Qy 596 TGGAAAGATGATGAACCTACCTCTGTATTTTTCAGACGGTCTGCACCTTCCAAGACGG 655
Db 9296 TACTTTGGAAGATTAAACGTTTACTTTGTTTTCAGAGCTGTCTGCAACTTTTACAGATGS 9355
Qy 656 TGCCCTGTGTCAGACCGAGCAATGGACCGGAAAGGAGACGACATAAACAAGAAAACTGAA 715
Db 9356 TGCAATGTTTCAGCATCAGGAGTGGGATGGGAAGGAAAGCACAATAACNAGAAAAATTGAA 9415
Qy 716 GGATGGGAGAGATGATCGTGGTGAGCATCAAGACCATGGCACCATCTGCGGATGGGCCCTG 775
Db 9416 AGATGGGAAATTTAGTGGTGAAGTCAAACTGCTGTGTCTCAGTCAGCTCTTGTGTG 9475
Qy 776 CAGGCACAGTTGTCTATAACCACTTCGGGTCTATGGTCTTTTAAACAAGAGAAAGAACTTA 835
Db 9476 CATTCATAGTTCAATAACTGTTCTATATCATTCATT---HACAGAACTCAGTTTG 9531
Qy 836 GGAGACCAATCTGAAAAATAACAAGTTAGAAACAGAGATCTCTCATTTGTCGAGCGAGCCCT 895
Db 9532 GAAAGAAAAAAGCAAAATAACAAGTTAAATACTAGAACACATAATATTATTAAGAAAGTCT 9591
Qy 896 TGTGGGCGGAGAGATGATGGG-----ATCCCAGATGTGGCTGCGAGCAGAGCCTG--- 947
Db 9592 AGTGGAGATAGAGAAAGTGCATGACATAGGAGAGGATTTGGCTGGAGTTGGGGGGAGTTC 9651
Qy 948 -----AGAGCTGGCAGGCCACCGAGCAGCCCTCTCCTGCT 982
Db 9652 TTGCTTTGCTGCCACGTCACTGAAATCTTGGCAAGCCACCACCAACTGCACATTTTTC 9711
Qy 983 ACATTTGATTTAAGTAAAGGATATTTGCCAAAAACAATGAATAATTTTAGAGATCATATCCA 1042
Db 9712 TATATCTTTTAAATAAAGGAATATTTGCCATAAACCAAAAGATAAAGATATATTTCAACACA 9771
Qy 1043 GTGCTTTTAGTCTGCGGGCAGCAATATACATATATAAACAACAGCAGCTCTAGGTCTTT- 1101
Db 9772 TTAGTTTTCTGTAAAAACATGTCATATTTAAGCAAAACCAACAGCTCTCTGCTCTCAAA 9831
Qy 1102 CTGAGTTTGAATCTCTGAGATGTGGTTTTCTGTAGTTGGTTTGAACAGCGTTTATAGGA 1161
Db 9832 CCYGTGAATCTGAAAGAAATTTTCTCATCTATGAGTAGATTTAGTGATTTCTGGGA 9891
Qy 1162 TTCTCCCC-----ACAACATGCTCTGAAATGTACAGTTGGCTGGAGACTC 1208
Db 9892 CTTTGATTTCTAATGTTTTTAATACCACCTGCTCTGGAATCTAAGGCTTAACCTAATCTTT 9951
Qy 1209 TATCTTTCTT-----CTCCTAGGAGTGTCTATGAACCAATGCCACTGCACTCGGGCTCA 1263
Db 9952 TTAATATCTTTCTCTTCTTAGGAGTGTCTATGAACCAATGTCACTGTACTCGGATCTA 10011
Qy 1264 TGAGAAAGGTGCAATGA 1279
Db 10012 TGAAAAAGTAGAATAA 10027
```

```
RESULT 3
US-09-724-676-2658
; Sequence 2658, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2658
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-2658

Query Match      17.1%; Score 218.4; DB 33; Length 838;
Best Local Similarity 61.2%; Pred. No. 3.5e-53;
Matches 427; Conservative 0; Mismatches 256; Indels 15; Gaps 4;

QY 130 AGGAGTGGAGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGCTGTATCATTTAC 189
Db 131 AGGAGTGGGAATAGCTTTTGGGAAAATGGCGCAATGGCCAGCCAGCTGTATCATTTAC 190
QY 190 GTGTGATGGCAACATCACCGTCAAAACCGAGAGCAGTGGAAGCAGCTGTGTCTC 249
Db 191 TTGTGATGGTAAACCTCACCATAAAACTGAGAGCACTTTGAAAACCAACACAGTTTC 250
QY 250 TTGTAACTGGGAGAGAGTTTGTGAAACGACAGCTGATGCGAGAAAACCTGGAGTCAG 309
Db 251 TTGTACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGCGAGAAAACCTCCGGTCCA 310
QY 310 CTACAACTACTGTGTAAGCGACAGAGCTTCTAGATTTTACAGATTAATTTGCAATTAACAA 369
Db 311 GTCGTGACCATGTTATGAATCACGAGCTTCTGAATGATAGGCTGTAT----CAATAAC 366
QY 370 TGTCTGTACTTACTGCCAAGGCTGACTGAAAAACTACTTTATGGAGTTGACTTTTGAT 429
Db 367 ATTTTACTGTTTATAGGCAGAACTTAATGAAAAAGTTATTTTAT--GAATTGAATTTTGTG 425
QY 430 AAATAGTAAAAAGTCCCGAGCTTAAGAAATGAAGACATCTTATGAGTTTCTAGATCGAAA 489
Db 426 AAATAGCAAAAGTATCAACTTCATAGAAATGGCAATCTTTTATTAGCTACTAGGTTG 485
QY 490 AGCACAATAGTTGTATTTGTGAACAAATCAGATATGATGGGTGGAGT----TCAGAGAGG 545
Db 486 AAACCCACAACTATTTGTAATAAATCAATATGGGTTAATGAAGTAGACTCAGAAAGGA 545
QY 546 AAAGCGAAGACTTGTGGAGT-----GGTGGGTCTCTGGGGTTCCTTCACTTTGA 599
Db 546 GAAGGTGAACAAATGTTGATTAAGGAGTTATGAGTCATGGAACCTCTTGTAAATGTACT 605
QY 600 AGATGATGAACCTAACCTACCTGTATTTTTCAGAGCGTCTGCACCTTCCAGACGGTGCC 659
Db 606 TGAAGATTAACAACTTTTACTTTGTTTTCAGAGCTGTCTGCACTTTACAGATGGTGA 665
QY 660 CTGGTCCAGCACCAATGGGACGGGAGGAGAGCAGATTAACAGAAAACTGAAGGAT 719
Db 666 TTGGTTTACGATCAGAGTGGGATGGGAGGAGGAGGAGCAATTAACAGAAAAATGAAGAT 725
QY 720 GGGAGATGATCGTGGTGAAGCATCAAGCAGCTGGCAGCATGTGGGATTTGGCCCTGCAGC 779
Db 726 GGGAAATAGTGGTGAAGTGTCAAACTGCTGTCTCAGTCAGTCAGCTTCTTGTGTGCAATT 785
QY 780 CACAGTTGTCAACCACTTCGGGTCAATTGGTTCTTTA 817
Db 786 CATAGTTTCATAACTGTTCTATATCATTTGATCAATTA 823
```

RESULT 5
US-09-724-676-2657
; Sequence 2657, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD

; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2657
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-2657

Query Match 15.7%; Score 201.4; DB 33; Length 1105;
Best Local Similarity 62.2%; Pred. No. 4.8e-48;
Matches 389; Conservative 0; Mismatches 221; Indels 15; Gaps 4;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTAAC 189
DB 131 AGGAGTGGGAATAGCTTTGCGAANAATGGCGCAATGGCCAGCCAGATTTGTATCATCAC 190
QY 190 GTGTGATGGCAACAACATCACCGGTCAAAACCCGAGAGCAGCTGAAGACGACTGTGTCTC 249
DB 191 TTGTGATGTTAAAAACCTCACCATAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTC 250
QY 250 TTGTAACCTGGGAGAGAACTTTGATGAACGACAGCTGATGGCAGAAAACTGAGGTGAG 309
DB 251 TTGTACCTGGGAGAGAACTTTGAGAAAAACCAAGCTGATGGCAGAAAACTCCGGTCCA 310
QY 310 CTACACATACCTGTGAAGCGACAGAACTTCTAGATTTTACAGATTAATAATGCAATTAACAA 369
DB 311 GTGCTGACCATGTTATGAATCACGAAGCTTCTGAATGATAGCTGTAT-CAATAAC 366
QY 370 TGTCTGTACTTACTGCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTTTGAT 429
DB 367 ATTTTACTGTTTATAGGCAGAACTTAATGAAAAAGTTATTTAT-GAATTTGAATTTTGC 425
QY 430 AAATTAGTAAAGTCCAGGACTAGAAATGAAGACATCTTATGAGTTTCTAGATCGAAA 489
DB 426 AAATTAGCAAAAGTATCAACTTTCATCATAAGATTTGCACTTTTATTTAGCTACTAGTTG 485
QY 490 AGCACAATGTTCTATTGTGAACAAATCAGTATGATGGGGTGGAGT---TCAGAGAGGG 545
DB 486 AAAACCAACAATTTTGTGAATAAATCAATATGGGTTAATGAAGTAGACTCAGAAAGGA 545
QY 546 AAAGCGAAGACTGTTGGAGT-----GGTGTGGGTCTCTGGGGTTCCTTCACTTTGGA 599
DB 546 GAAGGTGAAACAAATGTTGATTAAAGGAGTTATGAGTCATGGAACCTTTGTAATGTACT 605
QY 600 AGATGATGAACATACTACCTGTATTTTTCGACAGCGTCTGCACCTTCCAGACGGTGCC 659
DB 606 TCGAAGATTAAACCGTTTACTTTGTTTTCGACAGCTGTCTGCAACTTTACAGATGGTGA 665
QY 660 CTGGTCCAGCACCAATGGGACGGGAGGAGCAGCAGATTAACAGAAAACTGAAGGAT 719
DB 666 TTGGTTTCAGCATCAGAGTGGGATGGGAAGGAAGCAATTAACAGAAAAATTTGAAGAT 725
QY 720 GCGAAGATGATCGTGTGAGCATCA 744
DB 726 GCGAATATTAGTGGTGGAGTGTGCA 750

RESULT 6
US-09-724-676A-2657
; Sequence 2657, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2657

; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-2657
Query Match 15.7%; Score 201.4; DB 33; Length 1105;
Best Local Similarity 62.2%; Pred. No. 4.8e-48;
Matches 389; Conservative 0; Mismatches 221; Indels 15; Gaps 4;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTAAC 189
DB 131 AGGAGTGGGAATAGCTTTGCGAANAATGGCGCAATGGCCAGCCAGATTTGTATCATCAC 190
QY 190 GTGTGATGGCAACAACATCACCGGTCAAAACCCGAGAGCAGCTGAAGACGACTGTGTCTC 249
DB 191 TTGTGATGTTAAAAACCTCACCATAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTC 250
QY 250 TTGTAACCTGGGAGAGAACTTTGATGAACGACAGCTGATGGCAGAAAACTGAGGTGAG 309
DB 251 TTGTACCTGGGAGAGAACTTTGAGAAAAACCAAGCTGATGGCAGAAAACTCCGGTCCA 310
QY 310 CTACACATACCTGTGAAGCGACAGAACTTCTAGATTTTACAGATTAATAATGCAATTAACAA 369
DB 311 GTGCTGACCATGTTATGAATCACGAAGCTTCTGAATGATAGCTGTAT-CAATAAC 366
QY 370 TGTCTGTACTTACTGCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTTTGAT 429
DB 367 ATTTTACTGTTTATAGGCAGAACTTAATGAAAAAGTTATTTAT-GAATTTGAATTTTGC 425
QY 430 AAATTAGTAAAGTCCAGGACTAGAAATGAAGACATCTTATGAGTTTCTAGATCGAAA 489
DB 426 AAATTAGCAAAAGTATCAACTTTCATCATAAGATTTGCACTTTTATTTAGCTACTAGTTG 485
QY 490 AGCACAATGTTCTATTGTGAACAAATCAGTATGATGGGGTGGAGT---TCAGAGAGGG 545
DB 486 AAAACCAACAATTTTGTGAATAAATCAATATGGGTTAATGAAGTAGACTCAGAAAGGA 545
QY 546 AAAGCGAAGACTGTTGGAGT-----GGTGTGGGTCTCTGGGGTTCCTTCACTTTGGA 599
DB 546 GAAGGTGAAACAAATGTTGATTAAAGGAGTTATGAGTCATGGAACCTTTGTAATGTACT 605
QY 600 AGATGATGAACATACTACCTGTATTTTTCGACAGCGTCTGCACCTTCCAGACGGTGCC 659
DB 606 TCGAAGATTAAACCGTTTACTTTGTTTTCGACAGCTGTCTGCAACTTTACAGATGGTGA 665
QY 660 CTGGTCCAGCACCAATGGGACGGGAGGAGCAGCAGATTAACAGAAAACTGAAGGAT 719
DB 666 TTGGTTTCAGCATCAGAGTGGGATGGGAAGGAAGCAATTAACAGAAAAATTTGAAGAT 725
QY 720 GCGAAGATGATCGTGTGAGCATCA 744
DB 726 GCGAATATTAGTGGTGGAGTGTGCA 750

RESULT 7
US-09-724-676-2656
; Sequence 2656, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2656
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-2656
Query Match 15.7%; Score 201.4; DB 33; Length 1369;
Best Local Similarity 62.2%; Pred. No. 5.4e-48;

Matches 389; Conservative 0; Mismatches 221; Indels 15; Gaps 4;
QY 130 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
D 131 AGGAGTGGGAATAGCTTTGCGAAAAATGGCCAAATGGCCAGCCAGACTGTATCATTCAC 190
QY 190 GTGTGATGCGCAACAATCATCCGCTCAAAACCGAGAGCAGTGAAGACGACTGTCTCTC 249
D 191 TTGTGATGTTAAACCTCACCATAAAATCTGAGAGCACTTTGAAAAACAACACAGTTTTTC 250
QY 250 TTGTAACTCTGGAGAGAAATTTGATGAACACGACGCTGATGCGAGAAAAAAGCTGAGGTGAG 309
D 251 TTGTACCTCTGGAGAGAAATTTGAGAAACACACGCTGATGCGCAAAAAAAGCTCCGGTCCA 310
QY 310 CTACAACATACCTGTGAAGCGACAGAAAGCTTCTAGATTTTACAGATTTAAATTTGCATTTAAACAA 369
D 311 GTCTGACCATCTTATGAATCAGAACTCAGAGCTTCTGAATGATAGGCTGTAT----CAATAAC 366
QY 370 TGTCTGTACTTACTGCCAAGGCTGACTGAAAAAAGCTTTTATGGAGTTGACTTTTGTAT 429
D 367 ATTTTACTGTTTATAGGCGAAGCTTAAATGAAGAAAGTTATTTTAT--GAATTTGAATTTTGTCTC 425
QY 430 AAATAGTAAAGTCCCGAGGCTGAGAAATGAAGACATCTTATGAGTTTCTAGATCGAAA 489
D 426 AAATAGCAAAAGTATCACTTATCATAGATTTGCGCATCTTTTATAGCTACTAGGTTG 485
QY 490 AGCAGATAGTTGATTTGAGAAACAAATCATAGATTTGAGGTTGAGTAAATGAAGAAAGG 545
D 486 AAACCCACAACTATTTGGAATAAATCAATATGGTTAATGAAGTAGACTCAGAAAGGA 545
QY 546 AAAGCGAGAGCTTGTGGAGT-----GGTGTGGGCTCTGGGGTTCCTTCACTTTGGA 599
D 546 GAAGGTGAACAAATGTTGATTAGGAGGTTATGAGTCAATGAAAACTCTTGTATGTACT 605
QY 600 AGATGATGAATTAACCTGATTTTGGAGAGGCTGCGACCTTCCAGACGGTGCC 659
D 606 TGGAGATTTAAACGTTTACTTTGTTTGGCAGACTGCTGCACTTTACAGATGGTGA 665
QY 660 CTGGTCCAGCACAGCAATGGAGCGGAAGGAGGACGATTAACAGAAAACTGAAGAT 719
D 666 TTGGTTTCAGCATCAGGAGTGGGATGGGAGGAAAGCACAATAACAGAAAAATTTGAAGAT 725
QY 720 GGGAGATGATCGTGTGAGCATCA 744
D 726 GGGAAATTAGTGGTGGAGTGTGCA 750

RESULT 8
US-09-724-676A-2656
; Sequence 2656, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2656
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-2656

Query Match 15.7%; Score 201.4; DB 33; Length 1369;
Best Local Similarity 62.2%; Pred. No. 5.4e-48;
Matches 389; Conservative 0; Mismatches 221; Indels 15; Gaps 4;
QY 130 AGGAGTAGACTGGCTCTTAGGAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
D 131 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGGATTTGATCATCAC 190
QY 190 GTGTGATGCGCAACAACATCATCCGCTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249

Db 191 TTGTGATGGTAAAAAACCCTCACCATAAAAAATGAGAGCACTTTGAAAAACAACACAGTTTTTC 250
QY 250 TTGTAACTCTGGAGAGAAATTTGATGAACACGACGCTGATGCGAGAAAAAAGCTGAGGTGAG 309
D 251 TTGTACCTCTGGAGAGAAATTTGAGAAACACACGCTGATGCGCAAAAAAAGCTCCGGTCCA 310
QY 310 CTACAACATACCTGTGAAGCGACAGAAAGCTTCTAGATTTTACAGATTTAAATTTGCATTTAAACAA 369
D 311 GTCTGACCATCTTATGAATCAGAACTCAGAGCTTCTGAATGATAGGCTGTAT----CAATAAC 366
QY 370 TGTCTGTACTTACTGCCAAGGCTGACTGAAAAAAGCTTTTATGGAGTTGACTTTTGTAT 429
D 367 ATTTTACTGTTTATAGGCGAAGCTTAAATGAAGAAAGTTATTTTAT--GAATTTGAATTTTGTCTC 425
QY 430 AAATAGTAAAGTCCCGAGGCTGAGAAATGAAGACATCTTATGAGTTTCTAGATCGAAA 489
D 426 AAATAGCAAAAGTATCACTTATCATAGATTTGCGCATCTTTTATAGCTACTAGGTTG 485
QY 490 AGCAGATAGTTGATTTGAGAAACAAATCATAGATTTGAGGTTGAGTAAATGAAGTAGACTCAGAAAGGA 545
D 486 AAACCCACAACTATTTGGAATAAATCAATATGGTTAATGAAGTAGACTCAGAAAGGA 545
QY 546 AAAGCGAGAGCTTGTGGAGT-----GGTGTGGGCTCTGGGGTTCCTTCACTTTGGA 599
D 546 GAAGGTGAACAAATGTTGATTAGGAGGTTATGAGTCAATGAAAACTCTTGTATGTACT 605
QY 600 AGATGATGAATTAACCTGATTTTGGAGAGGCTGCGACCTTCCAGACGGTGCC 659
D 606 TGGAGATTTAAACGTTTACTTTGTTTGGCAGACTGCTGCACTTTACAGATGGTGA 665
QY 660 CTGGTCCAGCACAGCAATGGAGCGGAAGGAGGACGATTAACAGAAAACTGAAGAT 719
D 666 TTGGTTTCAGCATCAGGAGTGGGATGGGAGGAAAGCACAATAACAGAAAAATTTGAAGAT 725
QY 720 GGGAGATGATCGTGTGAGCATCA 744
D 726 GGGAAATTAGTGGTGGAGTGTGCA 750

RESULT 9
US-10-631-467-1237
; Sequence 1237, Application US/10631467
; GENERAL INFORMATION:
; APPLICANT: Genex Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1237
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-631-467-1237

Query Match 13.8%; Score 176.6; DB 58; Length 933;
Best Local Similarity 95.3%; Pred. No. 1.1e-40;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 130 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
D 112 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 171
QY 190 GTGTGATGCGCAACAACATCATCCGCTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
D 172 GTGTGATGCGCAACAACATCATCCGCTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 231

QY 250 TTGTAACCTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGGTGAC 309
DB 232 TTGTAACCTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGACGGT 291
QY 310 CTACACATAC 320
DB 292 CTGCACCTTCC 302

RESULT 10

US-10-764-420-1507
; Sequence 1507, Application US/10764420
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROS122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1507
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1507

Query Match 13.8%; Score 176.6; DB 62; Length 933;
Best Local Similarity 95.3%; Pred. No. 1.1e-40;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCAGACTGTATCATTTAC 189
DB 112 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCAGACTGTATCATTTAC 171
QY 190 GTGTGATGCGACACATCAGCTCAGGTCAAAACCGAGGACAGTGAAGACGACTGTCTTC 249
DB 172 GTGTGATGCGACACATCAGCTCAGGTCAAAACCGAGGACAGTGAAGACGACTGTCTTC 231
QY 250 TTGTAACCTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGGTGAC 309
DB 232 TTGTAACCTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGACGGT 291
QY 310 CTACACATAC 320
DB 292 CTGCACCTTCC 302

RESULT 11

US-10-029-386-5224
; Sequence 5224, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5224
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Homo sapiens

FEATURE:

; OTHER INFORMATION: MAP TO AF181449.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: EST HUMAN HIT: AA428498.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q01469, EVALUATE 1.00e-27
; OTHER INFORMATION: NT HIT: g113642859, EVALUATE 2.00e-93
US-10-029-386-5224

Query Match 13.7%; Score 175.6; DB 46; Length 537;
Best Local Similarity 67.6%; Pred. No. 1.6e-40;
Matches 277; Conservative 0; Mismatches 129; Indels 4; Gaps 2;
QY 126 CCACAGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCAGACTGTATCA 185
DB 98 CTACAGGAGTGGGAATAGCTTTGCCGAAAATATGGGCGCAATGGCCAAAGCAGATTGTATCA 157
QY 186 TTACGTGTATGGGCAACACATCAGGTCGTAACCGAGAGACACAGTGAAGACAGCTGTGT 245
DB 158 TCACCTGTGTATGGTAAACACCTCACCATAAAACCTTGAGAGCACTTTGAAAAACAACAGT 217
QY 246 TCTCTGTAACTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGG 305
DB 218 TTTCTGTACCTGGGAGAGAGTTTGTATGAAACGACAGCTGATGCGAGAGAAAACCTGAGG 277
QY 306 TCAGCTACACATCTGTGAAAGCGACAGAGCTTCTAGATTTTACAGATTTAAATTCATTCA 365
DB 278 TCAGTCTGTATGTTATGAAATCACAGAGCTTCTAGAAATGATAGGCTGTAT--CAAT 334
QY 366 ACAATGTCTGTACTTACTGCCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTT 425
DB 335 AACATTTTACTGTTTATAGGCAAGAACTTAATGAAAAAGTTATTTTAT-GAATTGAATTT 393
QY 426 TGATAAATTTAGTAAAGTCCCGAGGACTAAGAAATCAGACATCTTATAGTTTCTAGATC 485
DB 394 TGTCAATTTAGCAAAAGTATCACTTCATCATGAAATTTGGCACTCTTTATTAGTACTACTAG 453
QY 486 GAAAGACACATAGTTTGTATTTGTGAAACAAAATCAGTATGATGGGTGGAGT 535
DB 454 GTTGAACCAACCAAACTATTGTGATAAATAAAATCAATATGGTTAATGAGT 503

RESULT 12

US-09-704-424-17602
; Sequence 17602, Application US/09704424
; GENERAL INFORMATION:
; APPLICANT: Watson, James
; TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed
; TITLE OF INVENTION: by the Polynucleotides and Methods for Their Use
; FILE REFERENCE: 11000.1045U
; CURRENT APPLICATION NUMBER: US/09/704,424
; CURRENT FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 36881
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17602
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Mouse
US-09-704-424-17602

Query Match 13.7%; Score 175; DB 32; Length 323;
Best Local Similarity 94.8%; Pred. No. 1.8e-40;
Matches 181; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCAGACTGTATCATTTAC 189
DB 104 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCAGACTGTATCATTTAC 163
QY 190 GTGTGATGCGACACATCAGCTCAGGTCAAAACCGAGAGACAGTGAAGACGACTGTCTTC 249
DB 164 GTGTGATGCGACACATCAGCTCAGGTCAAAACCGAGAGACAGTGAAGACGACTGTCTTC 223

Qy	250	TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCGAGAAAAAAGCTGAGGTCAG	309
Db	224	TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCGAGAAAAAAGCTGAGACGGT	283
Qy	310	CTACAAACATAC	320
Db	284	CTGCACCTTCC	294
RESULT 13			
US-09-704-424-16807			
; Sequence 16807, Application US/09704424			
; GENERAL INFORMATION:			
; APPLICANT: Watson, James			
; TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed			
; TITLE OF INVENTION: by the Polynucleotides and Methods for Their Use			
; FILE REFERENCE: 11000.1045U			
; CURRENT APPLICATION NUMBER: US/09/704,424			
; CURRENT FILING DATE: 2000-11-01			
; NUMBER OF SEQ ID NOS: 36881			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 16807			
; LENGTH: 340			
; TYPE: DNA			
; ORGANISM: Mouse			
US-09-704-424-16807			
Query Match 13.7%; Score 175; DB 32; Length 340;			
Best Local Similarity 94.8%; Pred. No. 1.9e-40;			
Matches 181; Conservative 0; Mismatches 10; Indels 0; Gaps 0			
Qy	130	AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCAAGCCAGACTGTATCATTTAC	189
Db	117	AGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCAAGCCAGACTGTATCATTTAC	176
Qy	190	GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC	249
Db	177	GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC	236
Qy	250	TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCGAGAAAAAAGCTGAGGTCAG	309
Db	237	TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCGAGAAAAAAGCTGAGACGGT	296
Qy	310	CTACAAACATAC	320
Db	297	CTGCACCTTCC	307
RESULT 14			
US-09-704-424-21386			
; Sequence 21386, Application US/09704424			
; GENERAL INFORMATION:			
; APPLICANT: Watson, James			
; TITLE OF INVENTION: Polynucleotides, Polypeptides Expressed			
; TITLE OF INVENTION: by the Polynucleotides and Methods for Their Use			
; FILE REFERENCE: 11000.1045U			
; CURRENT APPLICATION NUMBER: US/09/704,424			
; CURRENT FILING DATE: 2000-11-01			
; NUMBER OF SEQ ID NOS: 36881			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 21386			
; LENGTH: 370			
; TYPE: DNA			
; ORGANISM: Mouse			
US-09-704-424-21386			
Query Match 13.7%; Score 175; DB 32; Length 370;			
Best Local Similarity 94.8%; Pred. No. 2e-40;			
Matches 181; Conservative 0; Mismatches 10; Indels 0; Gaps 0			
Qy	130	AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCAAGCCAGACTGTATCATTTAC	189

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 17:41:25 ; Search time 2543.5 Seconds
(without alignments)
2368.781 Million cell updates/sec

Title: US-09-788-074-2
Perfect score: 1279
Sequence: 1 atgggagcaactgtagc.....tctatgagaagtgcgaatga 1279

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16258531 seqs, 2355356875 residues

Total number of hits satisfying chosen parameters: 32517062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents NA New.*

- 1: /cgn2_6/prodata/2/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/prodata/2/pna/PCT_NEW_COMB.seq2.*
- 3: /cgn2_6/prodata/2/pna/US06_NEW_COMB.seq.*
- 4: /cgn2_6/prodata/2/pna/US07_NEW_COMB.seq.*
- 5: /cgn2_6/prodata/2/pna/US08_NEW_COMB.seq.*
- 6: /cgn2_6/prodata/2/pna/US09_NEW_COMB.seq.*
- 7: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq.*
- 8: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq2.*
- 9: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq3.*
- 10: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq4.*
- 11: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq5.*
- 12: /cgn2_6/prodata/2/pna/US10_NEW_COMB.seq6.*
- 13: /cgn2_6/prodata/2/pna/US11_NEW_COMB.seq.*
- 14: /cgn2_6/prodata/2/pna/US11_NEW_COMB.seq2.*
- 15: /cgn2_6/prodata/2/pna/US11_NEW_COMB.seq3.*
- 16: /cgn2_6/prodata/2/pna/US11_NEW_COMB.seq4.*
- 17: /cgn2_6/prodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142.6	11.1	664	1	PCT-US05-11532-965 Sequence 965, App
2	142.6	11.1	664	2	PCT-US04-42360-985 Sequence 985, App
3	127.6	10.0	1378	13	US-11-136-527-3136 Sequence 3136, App
4	125	9.8	420	17	US-60-680-544-19719 Sequence 19719, A
5	125	9.8	420	17	US-60-680-473-19719 Sequence 19719, A
6	125	9.8	662	14	US-11-051-454-102 Sequence 102, App
7	125	9.8	662	17	US-60-651-235-2921 Sequence 2921, App
8	125	9.8	672	17	US-60-651-235-2918 Sequence 2918, App
9	125	9.8	1083	17	US-60-651-235-2920 Sequence 2920, App
10	125	9.8	45762	11	US-10-940-774A-16651 Sequence 16651, A
11	119.2	9.3	453	8	US-10-450-763-13607 Sequence 13607, A
12	118.6	9.3	1286	17	US-60-651-235-2919 Sequence 2919, App
13	118.6	9.3	13286	11	US-10-990-328A-97919 Sequence 97919, A
14	118.6	9.3	207317	11	US-10-990-328A-95342 Sequence 95342, A
15	102.2	8.0	1569	8	US-10-450-763-13606 Sequence 13606, A
16	102.2	8.0	1569	8	US-10-450-763-25344 Sequence 25344, A

17	94	7.3	598	8	US-10-450-763-27568 Sequence 27568, A
18	78.8	6.2	2083	13	US-11-066-316A-9920 Sequence 9920, App
19	78.8	6.2	80581	13	US-11-066-316A-9928 Sequence 9928, App
20	78.4	6.1	201	13	US-11-066-316A-28733 Sequence 28733, A
21	78.4	6.1	201	13	US-11-066-316A-28739 Sequence 28739, A
22	76.4	6.0	8524	11	US-10-940-774A-13282 Sequence 13282, A
23	73	5.7	201	11	US-10-990-328A-331909 Sequence 331909, A
24	73	5.7	201	11	US-10-990-328A-533079 Sequence 533079, A
25	71.2	5.6	925	17	US-60-680-544-19272 Sequence 19272, A
26	71.2	5.6	925	17	US-60-680-544-19731 Sequence 19731, A
27	71.2	5.6	925	17	US-60-680-473-19272 Sequence 19272, A
28	71.2	5.6	925	17	US-60-680-473-19731 Sequence 19731, A
29	70	5.5	620	11	US-10-940-774A-1540 Sequence 1540, App
30	70	5.5	636	13	US-11-090-997-1125 Sequence 1125, App
31	70	5.5	816	17	US-60-680-544-19390 Sequence 19390, A
32	70	5.5	816	17	US-60-680-473-19390 Sequence 19390, A
33	69.8	5.5	201	13	US-11-066-316A-3574 Sequence 3574, App
34	69.8	5.5	804	13	US-11-066-316A-205 Sequence 205, App
35	69.6	5.4	628	1	PCT-US05-00517-121 Sequence 121, App
36	69.4	5.4	399	17	US-60-680-544-47783 Sequence 47783, A
37	69.4	5.4	399	17	US-60-680-473-47783 Sequence 47783, A
38	67.4	5.3	201	13	US-11-066-316A-3569 Sequence 3569, App
39	67.4	5.3	1464	13	US-11-066-316A-204 Sequence 204, App
40	66.8	5.2	1153	17	US-60-680-544-19120 Sequence 19120, A
41	66.8	5.2	1153	17	US-60-680-473-19120 Sequence 19120, A
42	65.6	5.1	601	11	US-10-940-774A-42044 Sequence 42044, A
43	65.6	5.1	8597	11	US-10-940-774A-12945 Sequence 12945, A
44	65.4	5.1	753	13	US-11-136-527-1596 Sequence 1596, App
45	65.2	5.1	16596	17	US-60-659-397-12005 Sequence 12005, A

ALIGNMENTS

RESULT 1

PCT-US05-11532-965
; Sequence 965, Application PC/TUS0511532
; GENERAL INFORMATION:
; APPLICANT: Porter, Mark
; APPLICANT: Higgs, Brandon
; APPLICANT: Mendrick, Donna
; APPLICANT: Elashoff, Michael
; TITLE OF INVENTION: HEPATOTOXICITY MOLECULAR MODELS
; FILE REFERENCE: 044921-5134-WO
; CURRENT APPLICATION NUMBER: PCT/US05/11532
; CURRENT FILING DATE: 2005-04-07
; PRIOR APPLICATION NUMBER: US 60/559,949
; PRIOR FILING DATE: 2004-04-07
; NUMBER OF SEQ ID NOS: 2005
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 965
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US05-11532-965

Query Match	11.1%;	Score 142.6;	DB 1;	Length 664;
Best Local Similarity	84.7%;	Pred. No. 7.8e-30;		
Matches	160;	Conservative	0;	Mismatches 29; Indels 0; Gaps 0;
QY	130	AGGAGTAGGAGCTGCTTTAGGAAGATGGCTGCCAGCAGCAGATGATCATTAC	189	
Db	119	AGGAGTAGGAGCTGCTTTAGGAAGATGGCTGCCAGCAGCAGATGATCATTAC	178	
QY	190	GTGTGATGTCACACATCAGCTCAAAACCCAGAGCAGAGTGAACACCACTGTGTCTC	249	
Db	179	CCTGCACACACAACTCCTCAAACTGAGACACAGTGAACACCACTGTGTCTC	238	
QY	250	TTGTAACTGGGAGAGAGTTTGTGAAACGACAGCTGATGGCAGAAAACCTGAGGTGAG	309	
Db	239	TTGCACCTGGGAGAGAGTTTGTGAAACGACAGCTGATGGCAGAAAACCTGAGGTGAG	298	
QY	310	CTACAACAT	318	

```
Db          299 CTGCACCTT 307
|||||
Query Match      11.1%; Score 142.6; DB 2; Length 664;
Best Local Similarity 84.7%; Pred. No. 7.8e-30;
Matches 160; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
Db 119 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 178

Qy 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAGACGACTGTGTCTTC 249
Db 179 CCTGCACACAAACACCTCACCGTCAAAACCGAGAGACAGTGAAGACGACTGTGTCTTC 238

Qy 250 TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAACCTGAGGTGAG 309
Db 239 TTGCACCTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAACCTGAGACGCT 298

Qy 310 CTACAACAT 318
Db 299 CTGCACCTT 307

RESULT 3
US-11-136-527-3136
; Sequence 3136, Application US/11136527
; GENERAL INFORMATION:
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3136
; LENGTH: 1378
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (335)..(335)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (856)..(856)
```

```
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-3136

Query Match      10.0%; Score 127.6; DB 13; Length 1378;
Best Local Similarity 62.4%; Pred. No. 1.8e-25;
Matches 118; Conservative 45; Mismatches 26; Indels 0; Gaps 0;

Qy 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
Db 315 WGGAGTGGGGYTGCGWCBWRNGAAGATGGWKKCATGSCCAACACGACTGCATCATTTWC 374

Qy 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAGACGACTGTGTCTTC 249
Db 375 CCTGWCRCVCAMSARSSKACCGSDMDWBKSAGSAGSACRGTGAWGACGACCGTCTTTTC 434

Qy 250 TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAACCTGAGGTGAG 309
Db 435 TTSCACCTTGRGAGAGAGWKKGATGAAACCAACACAGCTGATGGCAGGRMACTGWRGCGGT 494

Qy 310 CTACAACAT 318
Db 495 YTGCAWYTT 503

RESULT 4
US-60-680-544-19719
; Sequence 19719, Application US/60680544
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary
; FILE REFERENCE: 21590290000
; CURRENT APPLICATION NUMBER: US/60/680,544
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 19719
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Macaca Mulatta
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(420)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-544-19719

Query Match      9.8%; Score 125; DB 17; Length 420;
Best Local Similarity 78.8%; Pred. No. 6.5e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
Db 142 AGGAGTGGGAATAGCTTTGCGAAAAAATGGCGCAATGGCCAAAGCCAGACTGTATCATCAC 201

Qy 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAGACGACTGTGTCTTC 249
Db 202 TTGTATGGCAAAAACCTCACCATTAATAACTGAGAGACACTTTGAAAAACAACAGTTTC 261

Qy 250 TTGTAACTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAACCTGAGGTGAG 309
Db 262 TTGTACCTCTGGGAGAGAAATTTGAAGAAACCAACAGCTGATGGCAGAAACCTCAGACTGT 321

Qy 310 CTACAACAT 318
Db 322 CTGCAGCTT 330
```

```
RESULT 5
US-60-680-473-19719
; Sequence 19719, Application US/60680473
; GENERAL INFORMATION:
; APPLICANT: Cooper, Matthew
; APPLICANT: Kinch, Deborah
; APPLICANT: Rosenberg, Michael
; APPLICANT: Subramaniam, S. Sai
; APPLICANT: Szak, Suzanne
; APPLICANT: Li, Huo
; APPLICANT: Bandaru, Raj
; APPLICANT: Derbel, Maher
; TITLE OF INVENTION: Nucleotide Array Containing Polynucleotide Probes Complementary to
; FILE REFERENCE: Fragments of, Cynomolgus Monkey Genes and the Use Thereof
; CURRENT APPLICATION NUMBER: US/60/680,473
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 48714
; SOFTWARE: Patent Sequence Analysis Tool Version 1.0
; SEQ ID NO 19719
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Macaca Mulatta
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(420)
; OTHER INFORMATION: n = A,T,C or G
US-60-680-473-19719

Query Match          9.8%; Score 125; DB 17; Length 420;
Best Local Similarity 78.8%; Pred. No. 6.5e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCCAAGCCAGACTGTATCATTTAC 189
DB 142 AGGAGTGGGAATAGCTTTTGGGAATAATGGCGCAATGGCCAGCCAGACTGTATCATCAC 201
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 202 TTGTGATGGCAAAACCTCACCATATAAACTGAGAGCAGCTTTGAAAACAACACAGTTTC 261
QY 250 TTGTAACCTGGGAGAGAGTTTGTATGAACAGCAGCTGATGGCAGAAAACTGAGTCTAG 309
DB 262 TTGTACCTCTGGGAGAGAAATTTGAAGAAACCCAGCTGATGGCAGAAAACTCAGACTGT 321
QY 310 CTACACAT 318
DB 322 CTGACGCTT 330

RESULT 6
US-11-051-454-102
; Sequence 102, Application US/11051454
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; FILE REFERENCE: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/205,823

US-11-051-454-102
; Sequence 102, Application US/11051454
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsley, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; FILE REFERENCE: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/11/051,454
; CURRENT FILING DATE: 2005-02-04
; PRIOR APPLICATION NUMBER: US/10/205,823
```

```
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-051-454-102

Query Match          9.8%; Score 125; DB 14; Length 662;
Best Local Similarity 78.8%; Pred. No. 7.7e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCCAAGCCAGACTGTATCATTTAC 189
DB 126 AGGAGTGGGAATAGCTTTTCCGAAAATATGGCGCAATGGCCAGCCAGATTGTATCATCAC 185
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 186 TTGTGATGGTAAAAACCTCACCATATAAACTGAGAGCAGCTTTGAAAACAACACAGTTTC 245
QY 250 TTGTAACCTGGGAGAGAGTTTGTATGAACAGCAGCTGATGGCAGAAAACTGAGTCTAG 309
DB 246 TTGTACCTCTGGGAGAGAAAGTTTGAAGAAACCCAGCTGATGGCAGAAAACTCAGACTGT 305
QY 310 CTACACAT 318
DB 306 CTGCAACTT 314

RESULT 7
US-60-651-235-2921
; Sequence 2921, Application US/60651235
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001581
; CURRENT APPLICATION NUMBER: US/60/651,235
; CURRENT FILING DATE: 2005-02-10
; NUMBER OF SEQ ID NOS: 3512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2921
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-651-235-2921

Query Match          9.8%; Score 125; DB 17; Length 662;
Best Local Similarity 78.8%; Pred. No. 7.7e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCCAAGCCAGACTGTATCATTTAC 189
DB 126 AGGAGTGGGAATAGCTTTTCCGAAAATATGGCGCAATGGCCAGCCAGATTGTATCATCAC 185
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 186 TTGTGATGGTAAAAACCTCACCATATAAACTGAGAGCAGCTTTGAAAACAACACAGTTTC 245
QY 250 TTGTAACCTGGGAGAGAGTTTGTATGAACAGCAGCTGATGGCAGAAAACTGAGTCTAG 309
DB 246 TTGTACCTCTGGGAGAGAAAGTTTGAAGAAACCCAGCTGATGGCAGAAAACTCAGACTGT 305
```

QY 310 CTACAACAT 318
|||
Db 306 CTGCAACTT 314

RESULT 8

US-60-651-235-2918
; Sequence 2918, Application US/60651235
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001581
; CURRENT APPLICATION NUMBER: US/60/651,235
; CURRENT FILING DATE: 2005-02-10
; NUMBER OF SEQ ID NOS: 3512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2918
; LENGTH: 672
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-651-235-2918

Query Match 9.8%; Score 125; DB 17; Length 672;
Best Local Similarity 78.8%; Pred. No. 7.7e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTCAC 189
|||||
Db 136 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGACTGTATCATCAC 195
|||||
QY 190 GTGTGATGCGCAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTCTC 249
|||||
Db 196 TTGTGATGTTAAAACTTCCATATAAACTGAGAGCATTGTGAAGAACACACAGTTTTTC 255
|||||
QY 250 TTGTAACTCTGGAGAGAAAGTTTGTATGAACACGACACTGATGCGACAGAAACCTGAGGTCTAG 309
|||||
Db 256 TTGTACCTCTGGAGAGAAAGTTTGAAGAAACACAGCTGATGCGACAGAAACCTCAGACTGT 315
|||||
QY 310 CTACAACAT 318
|||
Db 316 CTGCAACTT 324

RESULT 9

US-60-651-235-2920
; Sequence 2920, Application US/60651235
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001581
; CURRENT APPLICATION NUMBER: US/60/651,235
; CURRENT FILING DATE: 2005-02-10
; NUMBER OF SEQ ID NOS: 3512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2920
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-651-235-2920

Query Match 9.8%; Score 125; DB 17; Length 1083;
Best Local Similarity 78.8%; Pred. No. 9.2e-25;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTCAC 189
|||||
Db 181 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGACTGTATCATCAC 240
|||||
QY 190 GTGTGATGCGCAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTCTC 249
|||||
Db 241 TTGTGATGTTAAAACTTCCATATAAACTGAGAGCATTGTGAAGAACACACAGTTTTTC 300
|||||
QY 250 TTGTAACTCTGGAGAGAAAGTTTGTATGAAGAACGACAGCTGATGCGACAGAAACCTGAGGTCTAG 309
|||||

Db 301 TTGTACCTCTGGAGAGAGTTTGAAGAAACACACAGCTGATGCGAGAAAACTCAGACTGT 360
|||||
QY 310 CTACAACAT 318
|||
Db 361 CTGCAACTT 369

RESULT 10

US-10-940-774A-16651/C
; Sequence 16651, Application US/10940774A
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774A
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16651
; LENGTH: 45762
; TYPE: DNA
; ORGANISM: Human
US-10-940-774A-16651

Query Match 9.8%; Score 125; DB 11; Length 45762;
Best Local Similarity 78.8%; Pred. No. 3.5e-24;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTCAC 189
|||||
Db 28038 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGACTGTATCATCAC 27979
|||||
QY 190 GTGTGATGCGCAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTCTC 249
|||||
Db 27978 TTGTGATGTTAAAACTTCCATATAAACTGAGAGCATTGTGAAGAACACACAGTTTTTC 27919
|||||
QY 250 TTGTAACTCTGGAGAGAAAGTTTGTATGAAGAACGACAGCTGATGCGAGAAAACTGAGGTCTAG 309
|||||
Db 27918 TTGTACCTCTGGAGAGAAAGTTTGAAGAAACACACAGCTGATGCGAGAAAACTCAGACTGT 27859
|||||
QY 310 CTACAACAT 318
|||
Db 27858 CTGCAACTT 27850

RESULT 11

US-10-450-763-13607
; Sequence 13607, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 13607
; LENGTH: 453
; TYPE: DNA

Query Match 9.3%; Score 118.6; DB 11; Length 207317;
Best Local Similarity 76.7%; Pred. No. 4e-22;
Matches 145; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
|||
Db 51254 AGGAGTGGGAATAGCTTTTCAGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATCATCAC 51195
|||
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGACAGTGAAGACGACTGTGTTCTC 249
|||
Db 51194 TTGTGATGGCAGAAACCTCACACAAAAACCGAGAGCACTTTGAAAAACAACACAGTTTTTC 51135
|||
QY 250 TTGTAACCTGGGAGAGAACTTTTGATGAACGACAGCTGATGGCAGAAAAAAGCTGAGGTGAG 309
|||
Db 51134 TTGTACCTGGGAGAGAGGATTTTGAAGAAACCAACAGCTGATGGCAGAAAAACAACAGACTGT 51075
|||
QY 310 CTACAACAT 318
|||
Db 51074 CTGCAACTT 51066
|||

RESULT 15

US-10-450-763-13606
; Sequence 13606, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 13606
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (236)..(2971)
; OTHER INFORMATION: 100% homologous to Homo sapiens protein kinase C mu, accession
; OTHER INFORMATION: number X75756, Smith-Waterman Score=4804.
US-10-450-763-13606

Query Match 8.0%; Score 102.2; DB 8; Length 1569;
Best Local Similarity 78.7%; Pred. No. 3.1e-18;
Matches 122; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 164 ATGGCCAGCCAGACTGTATCATTTAGTGGCAACAACATCACGGTCAAAACCGAG 223
|||
Db 1 ATGGCCAGCCAGATTGTATCATCTTGTATAGCAAAACCTCACCATAAAACTGAG 60
|||
QY 224 AGCAGAGTGAAGACGACTGTGTTCTTTGTAACCTGGGAGAGAAAGTTGATGAAACGACA 283
|||
Db 61 AGCACTTTGAAAAACAACACAGTTTTCTGTGTACCTCTGGGAGAGAAAGTTTGAGAAAAACA 120
|||
QY 284 GCTGATGGCAGAAAAAAGCTGAGTCACTACATCAT 318
|||
Db 121 GCTGATGGCAGAAAGAACTCAGACTGTTTGCACACTT 155
|||

Search completed: July 12, 2005, 22:16:13
Job time : 2545.5 secs

Db 61 GSCCAGTGGGATGATAAGGAATGAATCCTTGCTTATCATGTTACAAATTAAGCTCATTTTC 120
Qy 121 CATACCCACAGAGTAGAGCTGGCTCTTTAGGAAGATGGCTGCCATGGCCAGCCAGACTG 180
Db 121 CATACCCACAGAGTAGAGCTGGCTCTTTAGGAAGATGGCTGCCATGGCCAGCCAGACTG 180
Qy 181 TATCATTACGTGTGATGCAACAAATCAGCGTCAAAACCCAGAGACAGAGTGAAGACCAAC 240
Db 181 TATCATTACGTGTGATGCAACAAATCAGCGTCAAAACCCAGAGACAGAGTGAAGACCAAC 240
Qy 241 TGTGTTCTCTTGTAACTGGGAGAGAAGTTTGAATGAACGACAGCTGATGGCAGAAAAAC 300
Db 241 TGTGTTCTCTTGTAACTGGGAGAGAAGTTTGAATGAACGACAGCTGATGGCAGAAAAAC 300
Qy 301 TGAGTCACTGATCAACATCTGTGAAGCGACAGAAAGCTTTAGATTTACAGATTAATTTG 360
Db 301 TGAGTCACTGATCAACATCTGTGAAGCGACAGAAAGCTTTAGATTTACAGATTAATTTG 360
Qy 361 CATTAACAATGCTCTGACTTACTGCTCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTG 420
Db 361 CATTAACAATGCTCTGACTTACTGCTCAAGGGCTGACTGAAAAAACTACTTTATGGAGTTG 420
Qy 421 ACTTTTGATAAATTAGTAAAGTCCAGGACTTAAGAAATGAAGACATCTTATGAGTTTCT 480
Db 421 ACTTTTGATAAATTAGTAAAGTCCAGGACTTAAGAAATGAAGACATCTTATGAGTTTCT 480
Qy 481 AGATCGAAAAAGCACATAGTTGTTGTAACAAATCAGATGATGGGGTGGAGTTTCA 540
Db 481 AGATCGAAAAAGCACATAGTTGTTGTAACAAATCAGATGATGGGGTGGAGTTTCA 540
Qy 541 GAGGGAAGGCGAAGACACTTGTGGAGTGTGTGGGTCTCTGGGGTCTCTTCACTTTGGA 600
Db 541 GAGGGAAGGCGAAGACACTTGTGGAGTGTGTGGGTCTCTGGGGTCTCTTCACTTTGGA 600
Qy 601 GATGATGAACCTAACCTGCTGATTTTTCAGACGGTCTGCACCTTCCAAGACGGTGGCC 660
Db 601 GATGATGAACCTAACCTGCTGATTTTTCAGACGGTCTGCACCTTCCAAGACGGTGGCC 660
Qy 661 TGGTCCAGCACAGCAATGGGACGGGAGGAGACAGTACAGTAAAGAACTGAAGGATG 720
Db 661 TGGTCCAGCACAGCAATGGGACGGGAGGAGACAGTACAGTAAAGAACTGAAGGATG 720
Qy 721 GGAAGATGATCTGTGTAGCATCAAGCACTGGCACCCTGCTGGGATTTGGGCTCGAGCC 780
Db 721 GGAAGATGATCTGTGTAGCATCAAGCACTGGCACCCTGCTGGGATTTGGGCTCGAGCC 780
Qy 781 ACAGTTGTCTATACCACTTCGGGTCTATTTGTTCTTTTAAACAGAGAGGAAACTTAGGAG 840
Db 781 ACAGTTGTCTATACCACTTCGGGTCTATTTGTTCTTTTAAACAGAGAGGAAACTTAGGAG 840
Qy 841 ACAATCTGAAAAATTAACAGTTTAGAAACGAGAGTCTCTATTTGCTGAGGCGCCCTTTGG 900
Db 841 ACAATCTGAAAAATTAACAGTTTAGAAACGAGAGTCTCTATTTGCTGAGGCGCCCTTTGG 900
Qy 901 GGACGAGAGAGTGTGGGATCCAGATGTGGCTGCAGCAGAGCTGAGAGCTGGCAGGC 960
Db 901 GGACGAGAGAGTGTGGGATCCAGATGTGGCTGCAGCAGAGCTGAGAGCTGGCAGGC 960
Qy 961 CACCGAGCAGCCCTCTCTGGTACATTTGATTTAAGTAAAGGATTTTGGCAAAACACATG 1020
Db 961 CACCGAGCAGCCCTCTCTGGTACATTTGATTTAAGTAAAGGATTTTGGCAAAACACATG 1020
Qy 1021 AATAATTTAGATCATATCCAGTGTAGTCTGAGGCGCAAAATATACATATAAC 1080
Db 1021 AATAATTTAGATCATATCCAGTGTAGTCTGAGGCGCAAAATATACATATAAC 1080
Qy 1081 AAAACAGAGCTCTAGTCTTCTGATTTGAACTCTGAGATGTGGTTTTCTGTTAGGT 1140
Db 1081 AAAACAGAGCTCTAGTCTTCTGATTTGAACTCTGAGATGTGGTTTTCTGTTAGGT 1140
Qy 1141 TGGTTACAGCGTTTTATAGATTTCTGCCAACACATGCTCTGAAATGTACAGTTGGCC 1200
Db 1141 TGGTTACAGCGTTTTATAGATTTCTGCCAACACATGCTCTGAAATGTACAGTTGGCC 1200

Qy 1201 TGAGACTCTATCTTTCTTCTCTAGGAGTGTGATGAACAATGCCACTGCACCTCGGCT 1260
Db 1201 TGAGACTCTATCTTTCTTCTCTAGGAGTGTGATGAACAATGCCACTGCACCTCGGCT 1260
Qy 1261 CTATGAGAAGGTGCAATGA 1279
Db 1261 CTATGAGAAGGTGCAATGA 1279

RESULT 2

US-10-764-420-1507
; Sequence 1507, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1507
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1507

Query Match 13.8%; Score 176.6; DB 21; Length 933;
Best Local Similarity 95.3%; Pred. No. 4.2e-42;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 130 AGGAGTAGAGCTGGCTCTTAGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 189
Db 112 AGGAGTAGAGCTGGCTCTTAGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTTAC 171
Qy 190 GTGTGATGCGCAACACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
Db 172 GTGTGATGCGCAACACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 231
Qy 250 TTGTAACCTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAAACTGAGGTGAG 309
Db 232 TTGTAACCTGGGAGAGAAAGTTTGATGAAACGACAGCTGATGGCAGAAAACTGAGACGCT 291
Qy 310 CTACACATAC 320
Db 292 CTGCACCTTCC 302

RESULT 3

US-10-029-386-5224
; Sequence 5224, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5224
; LENGTH: 537

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF181449.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.5
; OTHER INFORMATION: SWISSPROT HIT: AA428498.1, EVALUATION 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: Q01469, EVALUATION 1.00e-27
; OTHER INFORMATION: NT HIT: g113642859, EVALUATION 2.00e-93
US-10-029-386-5224

Query Match 13.7%; Score 175.6; DB 16; Length 537;
Best Local Similarity 67.6%; Pred. No. 6.1e-42;
Matches 277; Conservative 0; Mismatches 129; Indels 4; Gaps 2;
QY 126 CCACGAGGTAGGACTGGCTCTTAGGAAGATGCTGCCATGGCCAGCCAGAGCTGATCA 185
DB 98 CTACAGGAGTGGGAATAGCTTTGCGAANAATGGCGCAATGGCCAGCAGATTTGATCA 157
QY 186 TTACGTGTGATGCAACAAACATCACCGTCAAAACCGAGAGCAGAGTGAAGAGCAGCTGTGT 245
DB 158 TCACCTGTGTGTAATAAACCTCACCTATAAACTGAGAGCACTTTGAAACAACACACT 217
QY 246 TCTTTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGG 305
DB 218 TTTCTTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGG 277
QY 306 TCAGCTACACATCTGCTGAAGCGCAGAGAGCTTCTAGATTTACAGATTAATTCATTA 365
DB 278 TCAGTGTGACATGTTATGAATATCAGAGAGCTTCTAGAAATGATAGGCTGTAT---CAAT 334
QY 366 ACAATCTGTACTTACTGCGAAGGCTGACTGAAAAAACTACTTTATGGAGTTGACTTT 425
DB 335 AACATTTTACTGTTATAGCGAAGAACTTAATGAAGAACTTATTTAT-GAATTGAATTT 393
QY 426 TGATAAATTAGTAAAGTCCCGAGACTAAGAAATGAAGACATCTTAAGAGTTTCTAGATC 485
DB 394 TGTCAAATTAGCAAAAGTATCAACTTCATCATAGAAATGGCATCTTTTATTAGCTACTAG 453
QY 486 GAAAGCACATAGTTCTTATGTCACAAATCAGTATGATGGGTGAGT 535
DB 454 GTTGAACACCAAACTATTGTGAATAAATCAATATGGTTAATGAAGT 503

RESULT 4
US-09-778-844-65
; Sequence 65, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
; APPLICANT: JOHANSEN, JEANETTE ELISABETH
; APPLICANT: SCHALLING, MARTIN
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
; FILE REFERENCE: 030307/0195
; CURRENT APPLICATION NUMBER: US/09/778,844
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 65
; LENGTH: 1623
; TYPE: DNA
; ORGANISM: Mus Musculus
US-09-778-844-65

Query Match 12.9%; Score 165; DB 9; Length 1623;
Best Local Similarity 92.1%; Pred. No. 1.8e-38;
Matches 174; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGAGCTGATCATTTAC 189
DB 612 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGAGCTGATCATTTAC 671

QY 190 GTGTGATGGCAACAAACATCACCGTCAAAACCGAGAGCAGTGAAGACGAGCTGTGTTCTC 249
DB 672 TTGTGATGGCAACAAACATCACCGTCAAAACCGTGAAGCAGTGAAGACGAGCTGTGTTCTC 731
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGGTCAG 309
DB 732 TTGTACCTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGGTCAG 791
QY 310 CTACAACAT 318
DB 792 CTGCACCTT 800

RESULT 5
US-10-316-253-202
; Sequence 202, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Peng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 202
; LENGTH: 695
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (59)..(466)
; OTHER INFORMATION:
US-10-316-253-202

Query Match 11.3%; Score 144.2; DB 16; Length 695;
Best Local Similarity 85.2%; Pred. No. 2.2e-32;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGAGCTGTATCATTTAC 189
DB 136 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGAGCTGTATCATTTAC 195
QY 190 GTGTGATGGCAACAAACATCACCGTCAAAACCGAGAGCAGTGAAGACGAGCTGTGTTCTC 249
DB 196 CCTGACGCGCAACACCTCACCGTCAAAACCTGAGAGCAGGTTGAAGACGAGCTGTGTTCTC 255
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGGTCAG 309
DB 256 TTGCACCTGGGAGAGAGTTTGTATGAACCAACAGCTGATGGCAGAAAACTGAGGTCAG 315
QY 310 CTACAACAT 318
DB 316 CTGCACCTT 324

RESULT 6
US-10-316-253-198
; Sequence 198, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Peng
; APPLICANT: Greis, Kenneth

```
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 198
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)..(451)
; OTHER INFORMATION:
US-10-316-253-198

Query Match      11.3%; Score 144.2; DB 16; Length 704;
Best Local Similarity 85.2%; Pred. No. 2.2e-32;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
DB 121 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 180
QY 190 GTGTATGGCAACAACATCAGCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 181 CCTGACGGCAACAACCTCACCGTCAAAACCTGAGAGCAGCGTGAAGACGACCGTGTCTC 240
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACTGTGTCTC 309
DB 241 TTGCACCTTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACCGTGTCTC 300

QY 310 CTACAACAT 318
DB 301 CTGCACCTT 309

RESULT 7
US-10-388-934-489
; Sequence 489, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boese, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 489
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-388-934-489

Query Match      11.3%; Score 144.2; DB 17; Length 704;
Best Local Similarity 85.2%; Pred. No. 2.2e-32;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
DB 121 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 180
QY 190 GTGTATGGCAACAACATCAGCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 181 CCTGACGGCAACAACCTCACCGTCAAAACCTGAGAGCAGCGTGAAGACGACCGTGTCTC 240

QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACTGTGTCTC 309
DB 241 TTGCACCTTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACCGTGTCTC 300

QY 310 CTACAACAT 318
DB 301 CTGCACCTT 309

RESULT 8
US-10-191-803-522
; Sequence 522, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 522
; LENGTH: 704
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 S69874
US-10-191-803-522

Query Match      11.3%; Score 144.2; DB 17; Length 704;
Best Local Similarity 85.2%; Pred. No. 2.2e-32;
Matches 161; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
DB 121 AGGAGTAGGCTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 180
QY 190 GTGTATGGCAACAACATCAGCGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
DB 181 CCTGACGGCAACAACCTCACCGTCAAAACCTGAGAGCAGCGTGAAGACGACCGTGTCTC 240
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACTGTGTCTC 309
DB 241 TTGCACCTTGGGAGAGAGTTTGTATGAACCGAGACAGTGAAGACGACCGTGTCTC 300

QY 310 CTACAACAT 318
DB 301 CTGCACCTT 309

RESULT 9
US-10-316-253-200
; Sequence 200, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
```


; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 43253
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-43253

Query Match 9.8%; Score 125; DB 18; Length 333;
Best Local Similarity 78.8%; Pred. No. 9e-27; Mismatches 40; Indels 0; Gaps 0;
Matches 149; Conservative 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCAAGCCAGCAGCTGTATCATTTAC 189
|||||
Db 73 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATCATCAC 132
|||||
QY 190 GTGTGATGGCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
|||||
Db 133 TTGTGATGGTAAAAACCTCACATAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTC 192
|||||
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAAACGACAGCTGATGGCAAAAACTGAGGTCTAG 309
|||||
Db 193 TTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGACTGT 252
|||||
QY 310 CTACAACAT 318
|||||
Db 253 CTGCAACTT 261
|||||

RESULT 13
US-09-804-014A-46
; Sequence 46, Application US/09804014A
; Publication No. US20030064489A1
; GENERAL INFORMATION:
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Vernet, Corine
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinketsu, Richard
; APPLICANT: Spaderna, Steven
; APPLICANT: Majumder, Kumud
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-721 US
; CURRENT APPLICATION NUMBER: US/09/804,014A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/188,316
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/188,277
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/189,139
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/189,140
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: 60/190,401
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/190,231
; PRIOR FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-014A-46

Query Match 9.8%; Score 125; DB 10; Length 406;
Best Local Similarity 78.8%; Pred. No. 1e-26;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCAAGCCAGCAGCTGTATCATTTAC 189

Db 76 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATCATCAC 135
|||||
QY 190 GTGTGATGGCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
|||||
Db 136 TTGTGATGGTAAAAACCTCACATAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTC 195
|||||
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAAACGACAGCTGATGGCAAAAACTGAGGTCTAG 309
|||||
Db 196 TTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGACTGT 255
|||||
QY 310 CTACAACAT 318
|||||
Db 256 CTGCAACTT 264
|||||

RESULT 14
US-10-242-535A-7126
; Sequence 7126, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7126
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-7126

Query Match 9.8%; Score 125; DB 17; Length 439;
Best Local Similarity 78.8%; Pred. No. 1.1e-26;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCAAGCCAGCAGCTGTATCATTTAC 189
|||||
Db 123 AGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATCATCAC 182
|||||
QY 190 GTGTGATGGCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACTGTGTCTC 249
|||||
Db 183 TTGTGATGGTAAAAACCTCACATAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTC 242
|||||
QY 250 TTGTAACTGGGAGAGAGTTTGTATGAAACGACAGCTGATGGCAAAAACTGAGGTCTAG 309
|||||
Db 243 TTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAACTCAGACTGT 302
|||||
QY 310 CTACAACAT 318
|||||
Db 303 CTGCAACTT 311
|||||

RESULT 15
US-10-085-783A-7126
; Sequence 7126, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002

```
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7126
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; US-10-085-783A-7126

Query Match      9.8%; Score 125; DB 18; Length 439;
Best Local Similarity 78.8%; Pred. No. 1.1e-26;
Matches 149; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 AGGAGTGGCAATAGCTTTTCGGAATAATGGCGCAATGGCCAAAGCCAGACTGTATCATCAC 182
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCCGAGAGCACAGTGAAGACGACTGTGTTCTC 249
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TTGTGATGGTAAAAAACCTCACCATAAAAACTGAGAGCACTTTGAAAAACAACACAGTTTTTC 242
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 TTGTAACCTGGGAGAGAGAGTTTGTATGAAACGACAGCTGATGGCAGAAAAAAGCTGAGTCTAG 309
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 TTGTACCTGGGAGAGAGAGTTTGTAGAAACCAACACAGCTGATGGCAGAAAAAAGCTGAGTCTGT 302
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 CTACAAACAT 318
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 CTGCAACTT 311
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: July 12, 2005, 19:05:19
Job time : 2257.89 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:34 ; Search time 418.141 Seconds
(without alignments)
9372.119 Million cell updates/sec

Title: US-09-788-074-4
Perfect score: 662
Sequence: 1 accgcgcagcagaccctc.....taataaatgtgtgtgct 662

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	662	100.0	662	4	AAS13247 Human DNA
2	662	100.0	662	6	ABV77978 Hypoxia-r
3	662	100.0	662	10	ADB75278 Prostate
4	662	100.0	662	10	ADH28827 Human chr
5	662	100.0	662	12	ADJ75120 Marker ge
6	662	100.0	662	12	ADN03852 Antipsori
7	662	100.0	662	12	ADO19263 Human PRO
8	662	100.0	662	12	ADP13321 Renal cel
9	662	100.0	662	13	ADR24747 Breast ca
10	662	100.0	662	13	ACN38822 Tumour-as
11	662	100.0	662	13	ADP54337 Human PRO
12	662	100.0	662	13	ADR52981 Drug ther
13	662	100.0	662	13	ADP25373 PRO polyp
14	662	100.0	662	13	ADR99018 Fatty aci
15	662	100.0	662	13	ADR66234 Human pro
16	662	100.0	662	13	ADR66576 Human pro
17	660	99.7	720	2	AAZ77538 Human ova
18	660	99.7	1071	10	ADB47413 Human cdn
19	660	99.7	1071	11	ADM86791 Human cdn
20	649	98.0	1072	6	AAS94888 Human DNA

21	625.6	94.5	681	13	ACN39645	Acn39645 Tumour-as
22	618.4	93.4	660	12	ADQ86530	Adq86530 Human tum
23	616.8	93.2	644	13	ACN39640	Acn39640 Tumour-as
24	590	89.1	615	6	ABK53828	Abk53828 Human hea
25	588	88.8	698	5	AA991764	AA991764 DNA encod
26	583	88.1	606	6	ABK53952	Abk53952 Human hea
27	543.2	82.1	627	9	ABK57420	Abk57420 Human fat
28	535.6	80.9	634	6	ABS59331	Abs59331 Human fat
29	535.6	80.9	634	10	ADJ38435	Adj38435 Human cdn
30	533.2	80.5	612	10	ADK11762	Adk11762 Breast ca
31	517	78.1	555	6	ABK53861	Abk53861 Human hea
32	512	77.3	519	6	ABK53994	Abk53994 Human hea
33	500	75.5	566	6	ABK53874	Abk53874 Human hea
34	481.4	72.7	540	9	AAU57419	Aau57419 Human fat
35	476.4	72.0	494	9	ACH35892	Ach35892 Human end
36	474.2	71.6	588	9	ACH27751	Ach27751 Human adu
37	460.8	69.6	491	6	ABK92039	Abk92039 DNA encod
38	451	68.1	461	6	ABK53980	Abk53980 Human hea
39	439.4	66.4	501	12	ACH70818	Ach70818 Human gen
40	429.8	64.9	479	9	ACH38376	Ach38376 Human end
41	408	61.6	408	2	AAQ66842	Aaq66842 Melanogen
42	406	61.3	435	8	ACD13195	AcD13195 cDNA enco
43	395.8	59.8	410	10	ADK11763	Adk11763 Breast ca
44	367	55.4	413	5	AAS13344	Aas13344 Human cdn
45	365.6	55.2	499	5	AAS13343	Aas13343 Human cdn

ALIGNMENTS

RESULT 1

AAS13247
ID AAS13247 standard; DNA; 662 BP.

AC AAS13247;

DT 18-DEC-2001 (first entry)

DE Human DNA encoding keratinocyte fatty acid binding protein, Mall.

KW Human; Mal 1; keratinocyte fatty acid binding protein; ds; anorectic;
KW antidiabetic; antilipemic; antiatherosclerotic; obesity; diabetes;
KW dyslipidaemia; atherosclerosis; antisense therapy.

OS Homo sapiens.

EH Key Location/Qualifiers

FT CDS 49..456

FT FT /tag= a

FT FT /product= "Mal 1"

FT FT 645..650

FT FT /tag= b

XX WO200160384-A1.

PN 23-AUG-2001.

PP 16-FEB-2001; 2001WO-US005019.

XX 17-FEB-2000; 2000US-0183106P.

XX (HARD) HARVARD COLLEGE.

XX Hotamisligil GS;

XX WPI; 2001-570550/64.

XX P-PSDB; AAU08674.

PT Reducing the level of circulating free fatty acids in a mammal, useful

PT for treating or preventing obesity, diabetes, dyslipidemia or

PT atherosclerosis, by administering a keratinocyte lipid binding protein

PT inhibitor.

XX

PS Disclosure; Page 2-3; 27pp; English.

XX The invention relates to reducing the level of circulating free fatty acids in a mammal comprising administering to a mammal a compound that inhibits expression (e.g. by antisense therapy) of Mal1 (keratinocyte fatty acid binding protein) or Mal1 activity. The method is useful for reducing the level of circulating free fatty acids in a mammal.

CC particularly in a mammal that is suffering from or at risk of developing obesity, diabetes, dyslipidaemia or atherosclerosis. The method is also useful for diagnosing these diseases. The present sequence encodes Human Mal 1

XX

XX Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 662; DB 4; Length 662;

Best Local Similarity 100.0%; Pred. No. 1.3e-164;

Mismatches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCCGACCCACCATGSCCAGATT 60

DB 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCCGACCCACCATGSCCAGATT 60

QY 61 CAGCAGCTGGAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120

DB 61 CAGCAGCTGGAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAAATGGCGCAATGGCCAGAGATTGTATC 180

DB 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAAATGGCGCAATGGCCAGAGATTGTATC 180

QY 181 ATCACTTGTGATGTTAAACCTCACCATAAAACTGAGCAGCTTTGAAACACACAG 240

DB 181 ATCACTTGTGATGTTAAACCTCACCATAAAACTGAGCAGCTTTGAAACACACAG 240

QY 241 TTTTCTGTACCTGGGAGAGAGTTTGAAGAAACACAGCTGTATGGCAGAAAACCTCAG 300

DB 241 TTTTCTGTACCTGGGAGAGAGTTTGAAGAAACACAGCTGTATGGCAGAAAACCTCAG 300

QY 301 ACTGTCTGCACTTTACAGATGGTGATTTGGTTACGATCAGGAGTGGATGGGAAGGA 360

DB 301 ACTGTCTGCACTTTACAGATGGTGATTTGGTTACGATCAGGAGTGGATGGGAAGGA 360

QY 361 AGCAACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGAGTGTCTATGAACAAT 420

DB 361 AGCAACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGAGTGTCTATGAACAAT 420

QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATTCATCATCTTTGGACAG 480

DB 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATTCATCATCTTTGGACAG 480

QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTTTCTTT 540

DB 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTTTCTTT 540

QY 541 CTTTCTTTTCTTACTGTGTCTCAATTTATCTTTATCATACATTTTACATGCAGCTAT 600

DB 541 CTTTCTTTTCTTACTGTGTCTCAATTTATCTTTATCATACATTTTACATGCAGCTAT 600

QY 601 TTTCAAGTGTGTGGATTAATTAGGATCATCTCTTTGGTTAATAAATGATGTTTGTG 660

DB 601 TTTCAAGTGTGTGGATTAATTAGGATCATCTCTTTGGTTAATAAATGATGTTTGTG 660

QY 661 CT 662

DB 661 CT 662

RESULT 2

ABV77978

ID ABV77978 standard; DNA; 662 BP.

XX

XX

AC

XX

DT 12-NOV-2002 (first entry)

XX Hypoxia-repressed protein coding sequence #46.

DE Cytostatic; vasotropic; tranquiliser; antiatherosclerotic; gene therapy; antiinflammatory; vulnary; gynecological; ophthalmological; vaccine;

KW hypoxia; tumorigenesis; angiogenesis; apoptosis; cancer;

KW ischaemic condition; reperfusion injury; retinopathy; neonatal stress;

KW preclapmsia; atherosclerosis; inflammatory condition; wound healing;

KW inflammation; erythropoiesis; hair loss; human; gene; ds.

XX

OS Homo sapiens.

XX

XX WO200246465-A2.

XX

PD 13-JUN-2002.

XX

XX 10-DEC-2001; 2001WO-GB005458.

XX

PR 08-DEC-2000; 2000GB-00030076.

PR 08-FEB-2001; 2001GB-00003156.

PR 25-OCT-2001; 2001GB-00025666.

XX

XX (OXFO-) OXFORD BIOMEDICA UK LTD.

XX

XX White J, Mundy CR, Ward NR, Krige D, Kingsman SM, Harris RA; Rayner WN;

PI

XX

DR WPI; 2002-627238/67.

XX

PT Identifying a gene involved in disease for treating hypoxia-regulated conditions, comprises comparing the transcriptome/proteome of two cell types under different conditions and identifying a differentially regulated gene.

PT

XX

PS Claim 23; Page 331; 538pp; English.

XX

CC The present invention relates to methods for identifying genes and proteins that are implicated in a specific disease or physiological condition. The method comprises comparing the transcriptome/proteome of a specialised cell type implicated in a disease or condition with that of a second specialised cell type, under two experimental conditions, and identifying a gene that is differentially regulated in the two specialised cell types under experimental conditions. ABV7873-ABV78116 and ABV65061-ABV65257 were identified using the methods of the invention.

CC The coding sequences and proteins are useful for treating a disease in a patient, for manufacture of a medicament for treating hypoxia-regulated conditions, and for regulating tumorigenesis, angiogenesis, apoptosis, biological response to hypoxia conditions, or hypoxic-associated pathology in a patient. The coding sequences and proteins are also useful for monitoring the therapeutic treatment of a disease or physiological condition, such as cancer, ischaemic conditions, reperfusion injury, retinopathy, neonatal stress, preclapmsia, atherosclerosis, inflammatory conditions, wound healing, inflammation, erythropoiesis or hair loss

XX

SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 6; Length 662;

Best Local Similarity 100.0%; Pred. No. 1.3e-164;

Mismatches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCCGACCCACCATGSCCAGATT 60

DB 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCCGACCCACCATGSCCAGATT 60

QY 61 CAGCAGCTGGAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120

DB 61 CAGCAGCTGGAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAAATGGCGCAATGGCCAGAGATTGTATC 180

DB 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAAATGGCGCAATGGCCAGAGATTGTATC 180

QY 181 ATCACTGTGATGGTAAACCTCACCATAAAACCTGAGGACATTTTGAAGAACACACAG 240
 DB |||||
 181 ATCACTGTGATGGTAAACCTCACCATAAAACCTGAGGACATTTTGAAGAACACACAG 240
 QY 241 TTTTCTTGACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAAAACTCAG 300
 DB |||||
 241 TTTTCTTGACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAAAACTCAG 300
 QY 301 ACTGTCGCACTTTACAGATGTCATTTGGTTTCAGCATCAGAGTGGGATGGGAAGAA 360
 DB |||||
 301 ACTGTCGCACTTTACAGATGTCATTTGGTTTCAGCATCAGAGTGGGATGGGAAGAA 360
 QY 361 AGCACAATACAGAAATTTGAAGATGGAAATTTAGTGGTGGTGTCTCATGACAT 420
 DB |||||
 361 AGCACAATACAGAAATTTGAAGATGGAAATTTAGTGGTGGTGTCTCATGACAT 420
 QY 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCACTTTGGACAG 480
 DB |||||
 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCACTTTGGACAG 480
 QY 481 GAGTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTGTTTCTTT 540
 DB |||||
 481 GAGTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTGTTTCTTT 540
 QY 541 CTTTCTTTTTCATTTACTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 DB |||||
 541 CTTTCTTTTTCATTTACTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 QY 601 TTTCAAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTGTG 660
 DB |||||
 601 TTTCAAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTGTG 660
 QY 661 CT 662
 DB ||
 661 CT 662

RESULT 3

ADB75278
 ID ADB75278 standard; cDNA; 662 BP.
 XX ADB75278;
 AC ADB75278;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Prostate cancer marker cDNA.
 XX
 KW Prostate; cancer; cytostatic; gene therapy; marker; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003009814-A2.
 XX
 PD 06-FEB-2003.
 XX
 PF 25-JUL-2002; 2002WO-US023913.
 XX
 PR 25-JUL-2001; 2001US-0307982P.
 XX
 PR 22-AUG-2001; 2001US-0314356P.
 PR
 PR 25-SEP-2001; 2001US-0325020P.
 PR
 PR 12-DEC-2001; 2001US-0341746P.
 PR
 PR 05-MAR-2002; 2002US-0362158P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
 PI Hoersht S, Kamatkar S, Wonsay AM, Glatt K, Zhao X, Anderson D;
 XX
 DR WPI; 2003-248033/24.
 XX
 PT New nucleic acid molecule, useful for diagnosing or treating prostate
 PT cancer.
 XX

PS Disclosure; SEQ ID NO 102; 99pp; English.
 XX
 CC The invention relates to newly discovered cancer markers associated with
 CC the cancerous state of prostate cells. Also disclosed is a method of
 CC assessing whether a patient is afflicted with prostate cancer. The method
 CC of the invention involves assessing whether a patient is afflicted with
 CC prostate cancer by comparing the level of expression of a marker in a
 CC patient sample and the normal level of expression of the marker in a
 CC control non-prostate cancer sample, where a significant increase in the
 CC level of expression of the marker in the patient sample and the normal
 CC level indicates that the patient is afflicted with prostate cancer.
 CC Nucleic acids of the invention are useful for diagnosing or treating
 CC prostate cancer, and may be useful in gene therapy. Sequences given in
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 10; Length 662;
 Best Local Similarity 100.0%; Pred. No. 1.3e-164;
 Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCGCGGAGCGACAGCCCTCTCTGACGCGCAGCCGCCGCCGCCACCCACCCAGCTT 60
 DB |||||
 1 ACCGCGGAGCGACAGCCCTCTCTGACGCGCAGCCGCCGCCGCCACCCACCCAGCTT 60
 QY 61 CAGCAGCTGGAAGGAGATGGCGCTTGGTGACAGCAAGGCTTTGATCAATACATGAAG 120
 DB |||||
 61 CAGCAGCTGGAAGGAGATGGCGCTTGGTGACAGCAAGGCTTTGATCAATACATGAAG 120
 QY 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAATTTGGCGCAATGGCCAGCAGATTGTATC 180
 DB |||||
 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAATTTGGCGCAATGGCCAGCAGATTGTATC 180
 QY 181 ATCACTGTGATGTTAAACCTCACCATAAAACCTGAGAGCAGCTTTGAAAACACACAG 240
 DB |||||
 181 ATCACTGTGATGTTAAACCTCACCATAAAACCTGAGAGCAGCTTTGAAAACACACAG 240
 QY 241 TTTTCTTGACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAAAACTCAG 300
 DB |||||
 241 TTTTCTTGACCTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAAAACTCAG 300
 QY 301 ACTGTCGCACTTTACAGATGTCATTTGGTTTCAGCATCAGAGTGGGATGGGAAGAA 360
 DB |||||
 301 ACTGTCGCACTTTACAGATGTCATTTGGTTTCAGCATCAGAGTGGGATGGGAAGAA 360
 QY 361 AGCACAATACAGAAATTTGAAGATGGAAATTTAGTGGTGGTGTCTCATGACAT 420
 DB |||||
 361 AGCACAATACAGAAATTTGAAGATGGAAATTTAGTGGTGGTGTCTCATGACAT 420
 QY 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCACTTTGGACAG 480
 DB |||||
 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAATTTCCATCATCACTTTGGACAG 480
 QY 481 GAGTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTGTTTCTTT 540
 DB |||||
 481 GAGTTAATTAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTGTTTCTTT 540
 QY 541 CTTTCTTTTTCATTTACTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 DB |||||
 541 CTTTCTTTTTCATTTACTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 QY 601 TTTCAAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTGTG 660
 DB |||||
 601 TTTCAAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTGTG 660
 QY 661 CT 662
 DB ||
 661 CT 662

```

RESULT 4
ADH28827 standard; DNA; 662 BP.
XX
AC
XX
DT 11-MAR-2004 (first entry)
XX
DE Human chronic myelogenous leukaemia (CML) gene marker #95.
XX
KW ds; chronic phase chronic myelogenous leukaemia; CP-CML;
KW blast crisis CML; BC-CML; human; chronic myelogenous leukaemia;
KW gene marker.
XX
OS Homo sapiens.
XX
PN US2003104426-A1.
XX
PD 05-JUN-2003.
XX
PF 14-JUN-2002; 2002US-00171581.
XX
PR 18-JUN-2001; 2001US-0298914P.
XX
PA (LINS//) LINSLEY P S.
PA (MAOM//) MAO M.
PA (DAIH//) DAI H.
PA (HEY//) HE Y.
PA (RADI//) RADICH J P.
XX
PI Linsley PS, Mao M, Dai H, He Y, Radich JP;
XX
XX WPI; 2003-787046/74.
XX
DR
XX
PT Classifying cell sample as chronic phase chronic myelogenous leukemia or
PT blast crisis chronic myelogenous leukemia by detecting difference in
PT expression of genes corresponding to the markers such as X15415, U89436.
XX
XX Disclosure; SEQ ID NO 95; 31pp; English.
XX
XX The invention relates to a method of classifying a cell sample as chronic
CC phase chronic myelogenous leukaemia (CP-CML) or blast crisis CML (BC-
CC CML). The method is useful for classifying a sample as CP-CML or BC-CML.
CC The present sequence represents a human chronic myelogenous leukaemia
CC (CML) gene marker used to distinguish blast crisis CML from chronic phase
CC CML.
XX
XX Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 662; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCATGGCCACAGTT 60
DB 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCATGGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGGACAGCAAGGCTTTGATGAATACATGAAG 120
DB 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGGACAGCAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGATGGGAATAGCTTTGGGAAAATGGCGCAATGGCAAGCCAGATTGTATC 180
DB 121 GAGCTAGGATGGGAATAGCTTTGGGAAAATGGCGCAATGGCAAGCCAGATTGTATC 180
QY 181 ATCACTTGTGATGTTAAACCTCCACATAAAATCTGAGAGCACCTTTGAAAAACAACAG 240
DB 181 ATCACTTGTGATGTTAAACCTCCACATAAAATCTGAGAGCACCTTTGAAAAACAACAG 240
QY 241 TTTTCTTTGACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCGAAAAAATCAG 300
DB 241 TTTTCTTTGACCTGGGAGAGAGTTTGAAGAAACACAGCTGATGGCGAAAAAATCAG 300

```

```

QY 301 ACTGCTGCAACTTTTACAGATGGTGCAATGTTGTTCCAGCATCAGGAGTGGATGGGAAGAA 360
DB |||||
QY 361 AGCAATAAACAGAAATTCGAAAGATGGGAAAATTTAGTGGTGGAGTGTGTATGAACAAT 420
DB |||||
QY 421 GTCACCTGTACTCGGATCTATGAAAAGTGAATAAAATTTCCATCATCTTTGGACAG 480
DB |||||
QY 481 CAGTTAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTTTCTTT 540
DB |||||
QY 541 CTTTCTTTTTCATTACTGTTTCAATTTATCTTATCATATAAACATTTTACATGCGCTAT 600
DB |||||
QY 601 TTCAAAGTGTGTGGATTAAATTAGGATCATCCCTTTGGTTAATAATAAATGTTTGTG 660
DB |||||
QY 661 CT 662
DB ||
QY 661 CT 662
DB |||||
RESULT 5
ADJ75120
ID ADJ75120 standard; DNA; 662 BP.
XX
AC ADJ75120;
XX
DT 20-MAY-2004 (first entry)
XX
DE Marker gene SEQ ID NO:372.
XX
KW bronchial asthma; chronic obstructive pulmonary disease;
KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW gene therapy; marker gene; gene; ds.
XX
OS Homo sapiens.
XX
PN EP1394274-A2.
XX
PD 03-MAR-2004.
XX
PF 04-AUG-2003; 2003EP-00254857.
XX
PR 06-AUG-2002; 2002JP-00229312.
PR 20-MAR-2003; 2003JP-00077212.
XX
PA (GENO-) GENOX RES INC.
XX
PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuvara K;
XX
XX WPI; 2004-193155/19.
XX
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample
PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
XX
PS Claim 1; SEQ ID NO 372; 241pp; English.
XX
XX The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic

```

obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

XX
SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 12; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGAGGAGAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCCATGGCCACAGTT 60
DB |||||
QY 1 ACCGCGGAGGAGAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCCATGGCCACAGTT 60
DB |||||

QY 61 CAGCAGCTGGAAGAGAGTGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
DB |||||

QY 61 CAGCAGCTGGAAGAGAGTGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
DB |||||

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
DB |||||

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
DB |||||

QY 181 ATCACTTGTGATGGTAAACCTCAACATAAAACCTGAGAGCAGCTTTTGAACCAACACAG 240
DB |||||

QY 181 ATCACTTGTGATGGTAAACCTCAACATAAAACCTGAGAGCAGCTTTTGAACCAACACAG 240
DB |||||

QY 241 TTTTCTTGTACCTCGGAGAGAGTTTGAAGAAACCCACAGCTGTGGCGAAGAACTCAG 300
DB |||||

QY 241 TTTTCTTGTACCTCGGAGAGAGTTTGAAGAAACCCACAGCTGTGGCGAAGAACTCAG 300
DB |||||

QY 301 ACTGCTGCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGAGTGGATGGGAAGAA 360
DB |||||

QY 301 ACTGCTGCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGAGTGGATGGGAAGAA 360
DB |||||

QY 361 AGCAACAATAAGAAAAATTGAAAGATGGGAAATTAGTGGTGGAGTGTGTCATGAACAAT 420
DB |||||

QY 361 AGCAACAATAAGAAAAATTGAAAGATGGGAAATTAGTGGTGGAGTGTGTCATGAACAAT 420
DB |||||

QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATTCATCATCATCTTGGACAG 480
DB |||||

QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATTCATCATCATCTTGGACAG 480
DB |||||

QY 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTTTT 540
DB |||||

QY 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTTTT 540
DB |||||

QY 541 CTTTCTTTTTCATTAAGTGTTCATTAATCTTTATCAATAAATTTTACATGAGCTAT 600
DB |||||

QY 541 CTTTCTTTTTCATTAAGTGTTCATTAATCTTTATCAATAAATTTTACATGAGCTAT 600
DB |||||

QY 601 TTCAAGTGTGTGATTAATAGGATCATCCCTTTGGTTTAAATAAATGTTTGTG 660
DB |||||

QY 601 TTCAAGTGTGTGATTAATAGGATCATCCCTTTGGTTTAAATAAATGTTTGTG 660
DB |||||

QY 661 CT 662
DB ||
661 CT 662

RESULT 6
ADN03852
ID ADN03852 standard; cDNA; 662 BP.
XX
AC ADN03852;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #125.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
DR WPI; 2004-305105/28.
DR P-PSDB; ADN03853.
XX
XX New PRO nucleic acid or polypeptide, useful for preparing a
XX pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
XX
XX Claim 1; SEQ ID NO 246; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
XX treating psoriasis or a sequence having at least 80% identity to the
XX above sequences. The nucleic acid is useful for preparing a composition
XX for diagnosing or treating psoriasis in a mammal. This sequence
XX corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 12; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGAGGAGAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCCATGGCCACAGTT 60
DB |||||
QY 1 ACCGCGGAGGAGAGCCCTCTCTGACGCGCAGCCGCGCCGACCCACCCATGGCCACAGTT 60
DB |||||

QY 61 CAGCAGCTGGAAGAGAGTGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
DB |||||

QY 61 CAGCAGCTGGAAGAGAGTGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
DB |||||

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
DB |||||

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
DB |||||

QY 181 ATCACTTGTGATGGTAAACCTCAACATAAAACCTGAGAGCAGCTTTTGAACCAACACAG 240
DB |||||

QY 181 ATCACTTGTGATGGTAAACCTCAACATAAAACCTGAGAGCAGCTTTTGAACCAACACAG 240
DB |||||

QY 241 TTTTCTTGTACCTCGGAGAGAGTTTGAAGAAACCCACAGCTGTGGCGAAGAACTCAG 300
DB |||||

QY 241 TTTTCTTGTACCTCGGAGAGAGTTTGAAGAAACCCACAGCTGTGGCGAAGAACTCAG 300
DB |||||

QY 301 ACTGCTGCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGACTGGGATGGGAAGGA 360
 Db |||||
 QY 301 ACTGCTGCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGACTGGGATGGGAAGGA 360
 Db |||||
 QY 361 AGCACAATAACAAGAAAATTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420
 Db |||||
 QY 361 AGCACAATAACAAGAAAATTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420
 Db |||||
 QY 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAAATTCATCATCACTTTGGACAG 480
 Db |||||
 QY 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAAATTCATCATCACTTTGGACAG 480
 Db |||||
 QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATACGTGTTCTTT 540
 Db |||||
 QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATACGTGTTCTTT 540
 Db |||||
 QY 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 Db |||||
 QY 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600
 Db |||||
 QY 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
 Db |||||
 QY 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
 Db |||||
 QY 661 CT 662
 Db |||||
 QY 661 CT 662

RESULT 7

ADP19263

ID ADP19263 standard; cDNA; 662 BP.

AC

ADP19263;

DT

12-AUG-2004 (first entry)

DE

Human PRO polynucleotide #98.

KW

Human; PRO; gene; ss; immune related disorder;

KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;

KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;

KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;

KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;

KW renal disease; demyelinating disease; central nervous system;

KW peripheral nervous system; demyelinating polyneuropathy;

KW Guillain-Barre syndrome;

KW chronic inflammatory demyelinating polyneuropathy.

XX

OS Homo sapiens.

XX

PN WO2004043361-A2.

XX

PD 27-MAY-2004.

XX

PF 06-NOV-2003; 2003WO-US035268.

XX

PR 08-NOV-2002; 2002US-0425235P.

XX

XX (GETH) GENENTECH INC.

PA

PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;

PI Wood WI, Wu ID;

XX

DR WPI; 2004-420067/39.

XX

DR P-PSDB; ADP19264.

XX

PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for

PT treating an immune related disorder such as systemic lupus erythematosus,

PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or

PT spondyloarthritis.

XX

PS Claim 1; SEQ ID NO 204; 1731pp; English.

XX The invention relates to human PRO polypeptides and the polynucleotides
 CC encoding them. The polypeptides and polynucleotides are useful for
 CC treating and diagnosing immune related disorders in mammals. The immune
 CC related disorders include systemic lupus erythematosus, rheumatoid
 CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
 CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
 CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
 CC mellitus, immune-mediated renal disease, demyelinating diseases of the
 CC central or peripheral nervous system, demyelinating polyneuropathy,
 CC Guillain-Barre syndrome and chronic inflammatory demyelinating
 CC polyneuropathy. This sequence represents a human PRO polynucleotide of
 CC the invention.

SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 12; Length 662;

Best Local Similarity 100.0%; Pred. No. 1.3e-164;

Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGACGACAGACCCCTCTCTGACCGCCAGCCGCGCCGACCCACCATGGCCACAGTT 60

Db 1 ACCGCGGACGACAGACCCCTCTCTGACCGCCAGCCGCGCCGACCCACCATGGCCACAGTT 60

QY 61 CAGCAGCTGGGAAGGAGATGGCGCTGTGCACAGCAAGGCTTTGATGAATACATGAAG 120

Db 61 CAGCAGCTGGGAAGGAGATGGCGCTGTGCACAGCAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAATGGCGCAATGGCCAGGAGATTGTATC 180

Db 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAATGGCGCAATGGCCAGGAGATTGTATC 180

QY 181 ATCACTTGTGATGTAAAAAATCTCACCATAAAATGAGAGACACTTTGAAACACACAG 240

Db 181 ATCACTTGTGATGTAAAAAATCTCACCATAAAATGAGAGACACTTTGAAACACACAG 240

QY 241 TTTTCTGTACCTGGGAGAGAGTTTGAAGAAAACACAGCTGTATGGCAGAAAAAATCTCAG 300

Db 241 TTTTCTGTACCTGGGAGAGAGTTTGAAGAAAACACAGCTGTATGGCAGAAAAAATCTCAG 300

QY 301 ACTGCTCCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGAGTGGGATGGGAAGGA 360

Db 301 ACTGCTCCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGAGTGGGATGGGAAGGA 360

QY 361 AGCACAATAACAAGAAAATGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420

Db 361 AGCACAATAACAAGAAAATGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420

QY 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAAATTCATCATCACTTTGGACAG 480

Db 421 GTCACCTGTACTCGGATCTATGAAAAGTAGAATAAAAATTCATCATCACTTTGGACAG 480

QY 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATACGTGTTCTTT 540

Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATACGTGTTCTTT 540

QY 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600

Db 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCAATAAATTTTACATGCACTAT 600

QY 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660

Db 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660

QY 661 CT 662

Db |||||

QY 661 CT 662

Db |||||

RESULT 8

ADP13321

ID ADP13321 standard; DNA; 662 BP.

XX ADP13321;
 AC 26-AUG-2004 (first entry)
 DT XX Renal cell carcinoma differentially expressed gene #57.
 DE ds; diagnosis; non-blood disease; solid tumor; gene expression;
 KW peripheral blood mononuclear cell; renal cell carcinoma; prostate cancer;
 KW head/neck cancer; differential expression.
 XX Homo sapiens.
 OS WO2004048933-A2.
 PN 10-JUN-2004.
 PD XX
 PP 21-NOV-2003; 2003WO-US037481.
 XX 21-NOV-2002; 2002US-0427982P.
 PR 03-APR-2003; 2003US-0459782P.
 XX (AMHP) WYETH.
 PA (TWIN/) TWINE N C.
 PA (BURC/) BURCZYNSKI M E.
 PA (TREP/) TREPICCHIO W L.
 PA (DORN/) DORNER A.
 PA (STOV/) STOVER J A.
 PA (SLOW/) SLOWI D K.
 XX Twine NC, Burczynski ME, Trepicchio WL, Dorner A, Stover JA;
 PI Sloni DK;
 PI WPI; 2004-460799/43.
 DR XX
 XX Diagnosing non-blood disease such as solid tumor, involves comparing
 PT differential expression profile of specific genes in peripheral blood
 PT sample of subject with reference expression profile of specific genes.
 XX Disclosure; SEQ ID NO 57; 350pp; English.
 XX The invention relate to a method of diagnosing (M1) non-blood disease
 CC such as solid tumor by providing peripheral blood sample of human having
 CC non-blood disease, and comparing an expression profile of specific genes
 CC in the peripheral blood sample to reference expression profile of the
 CC genes, where each of the genes is differentially expressed in peripheral
 CC blood mononuclear cells (PBMCs) of patients having the disease as
 CC compared to PBMCs of normal humans. The method is useful for diagnosing
 CC non-blood disease such as solid tumor. The solid tumor is chosen from
 CC renal cell carcinoma (RCC), prostate cancer and head/neck cancer. The
 CC peripheral blood sample comprises enriched PBMCs. The peripheral blood
 CC sample is a whole blood sample (claimed). (M1) is useful for identifying
 CC genes that are differentially expressed in peripheral blood samples
 CC isolated at different stages of progression, development or treatment of
 CC RCC and/or other solid tumors. This sequence corresponds to a gene that
 CC is differentially expressed and detected by the method of the invention.
 CC (Note: this sequence is not given as part of the printed specification
 CC but was obtained from WIPO in electronic format at
 CC ftp.wipo./pub/published_pct_sequences).
 XX Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 662; DB 12; Length 662;
 Best Local Similarity 100.0%; Pred. No. 1.3e-164;
 Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCGCCGACGACACCCCTCTCTGACGGCAGCCCGCCGACCCACCATGGCCACAGTT 60
 DB 1 ACCGCCGACGACACCCCTCTCTGACGGCAGCCCGCCGACCCACCATGGCCACAGTT 60
 QY 61 CAGCAGCTGGAGGAGATGGCGCTCTGGCAGCAAGAGGCTTGTGATGAATACATGAG 120
 DB 61 CAGCAGCTGGAGGAGATGGCGCTCTGGCAGCAAGAGGCTTGTGATGAATACATGAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTTGGCAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
 DB 121 GAGCTAGGAGTGGGAATAGCTTTTGGCAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
 QY 181 ATCAGCTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTTGA AAAACAACACAG 240
 DB 181 ATCAGCTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTTGA AAAACAACACAG 240
 QY 241 TTTTCTTGTACCTCGGAGAGAGTGTGAAAGAAACACACAGCTGATGCGCAGAAAACCTCAG 300
 DB 241 TTTTCTTGTACCTCGGAGAGAGTGTGAAAGAAACACACAGCTGATGCGCAGAAAACCTCAG 300
 QY 301 ACTGCTGCAACTTTTACAGATGGTGGTTCAGCTCAGGAGTGGGATGGGAAGGAA 360
 DB 301 ACTGCTGCAACTTTTACAGATGGTGGTTCAGCTCAGGAGTGGGATGGGAAGGAA 360
 QY 361 AGCACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420
 DB 361 AGCACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTATGAACAAT 420
 QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCTTGGACAG 480
 DB 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCTTGGACAG 480
 QY 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAGTCAATGAGCAAAATCTCCATCTGTTCTTT 540
 DB 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAGTTCAGTCAATGAGCAAAATCTCCATCTGTTCTTT 540
 QY 541 CTTTCTTTTTCATCTGTTTCAATATCTTATCATATAAACATTTTACATGCAGCTAT 600
 DB 541 CTTTCTTTTTCATCTGTTTCAATATCTTATCATATAAACATTTTACATGCAGCTAT 600
 QY 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 DB 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAATGTTGTG 660
 QY 661 CT 662
 DB 661 CT 662

RESULT 9
 ADR24747
 ID ADR24747 standard; DNA; 662 BP.
 XX ADR24747;
 XX DT 21-OCT-2004 (first entry)
 XX DE Breast cancer prognosis marker #608.
 XX ds; breast cancer; prognosis; gene expression; diagnosis.
 XX Homo sapiens.
 XX WO2004065545-A2.
 XX 05-AUG-2004.
 XX 15-JAN-2004; 2004WO-US001100.
 XX 15-JAN-2003; 2003US-00342887.
 XX (ROSE-) ROSETTA INPHARMATICS LLC.
 XX (NECA-) NETHERLANDS CANCER INST.
 XX Van't Veer LJ, He Y;
 XX WPI; 2004-593473/57.
 XX Classifying a breast cancer patient according to prognosis comprises
 PT determining the similarity between the level of expression of each of

PT five genes in a cell sample taken from patient, to control levels.
XX Disclosure; SEQ ID NO 608; 226pp; English.
PS
XX
XX
CC The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
XX

Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 13; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCCGACGAGACCCCTCTCTGACCGCCGCGCCGCGCACCCACCATGGCCACAGTT 60
Db 1 ACCGCCGACGAGACCCCTCTCTGACCGCCGCGCCGCGCACCCACCATGGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAGATGGCGCTGTCGACGACCAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTGTCGACGACCAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAATGGCGCAATGGCCCAAGCTTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTGGCAAAATGGCGCAATGGCCCAAGCTTGTATC 180
QY 181 ATCACTTGTGATGTTAAACCTCAACATAAACTGAGAGCATTGTAACACACAG 240
Db 181 ATCACTTGTGATGTTAAACCTCAACATAAACTGAGAGCATTGTAACACACAG 240
QY 241 TTTTCTTCTACCTGGGAGAGAGTTTGAAGAAACACAGCTGTATGGCAGAAACTCAG 300
Db 241 TTTTCTTCTACCTGGGAGAGAGTTTGAAGAAACACAGCTGTATGGCAGAAACTCAG 300
QY 301 ACTGCTGCACTTTACAGATGGTGCTTTCAGCATCAGGATGGGATGGGAAGAA 360
Db 301 ACTGCTGCACTTTACAGATGGTGCTTTCAGCATCAGGATGGGATGGGAAGAA 360
QY 361 AGCAATAACAAAGAAATTTGAAAGATGGGAAATTTAGTGGAGTGTCTATGAACAT 420
Db 361 AGCAATAACAAAGAAATTTGAAAGATGGGAAATTTAGTGGAGTGTCTATGAACAT 420
QY 421 GTCACCTGTACTCGATCTATGAAAGATGAGTAATAAATTCATCATCTTTGGACAG 480
Db 421 GTCACCTGTACTCGATCTATGAAAGATGAGTAATAAATTCATCATCTTTGGACAG 480
QY 481 GAGTTAATTAAGAGATACCAAGCTCAGTTCAATGAGCAATCTCCATAGTTTCTTT 540
Db 481 GAGTTAATTAAGAGATACCAAGCTCAGTTCAATGAGCAATCTCCATAGTTTCTTT 540
QY 541 CTTTCTTTTTCATCTGTTCAATTTATCTTTATCAATAAATTTTACATGAGCTAT 600
Db 541 CTTTCTTTTTCATCTGTTCAATTTATCTTTATCAATAAATTTTACATGAGCTAT 600
QY 601 TTCAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTCTG 660
Db 601 TTCAAGTGTGTTGATTAATAGGATCATCCCTTTGGTTAATAAATGTTTCTG 660
QY 661 CT 662
Db 661 CT 662

RESULT 10
ACN38822
ID ACN38822 standard; cdna; 662 BP.

XX ACN38822;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325115, SEQ ID NO:2575.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
FN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX P-PSDB; ABM80997.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 2575; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;
Query Match 100.0%; Score 662; DB 13; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCCGACGAGACCCCTCTCTGACCGCCGCGCCGCGCACCCACCATGGCCACAGTT 60
Db 1 ACCGCCGACGAGACCCCTCTCTGACCGCCGCGCCGCGCACCCACCATGGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAGATGGCGCTGTCGACGACCAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTGTCGACGACCAAGGCTTTGATGAATACATGAAG 120

QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATCCATCATCAGCTTGGACAG 480
Db 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAATCCATCATCAGCTTGGACAG 480
QY 481 GAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTT 540
Db 481 GAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTT 540
QY 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCATATAACATTTTACATGAGCTAT 600
Db 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCATATAACATTTTACATGAGCTAT 600
QY 601 TTCAAAGTGTGTTGATTAAATTAGATCATCCCTTTGGTTAATAATAAATGTTTGTG 660
Db 601 TTCAAAGTGTGTTGATTAAATTAGATCATCCCTTTGGTTAATAATAAATGTTTGTG 660
QY 661 CT 662
Db 661 CT 662

RESULT 12
ID ADR52981 standard; DNA; 662 BP.
XX ADR52981;
AC ADR52981;
DT 18-NOV-2004 (first entry)
XX Drug therapy altered expressed gene #332.
DE drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; ds.
XX Homo sapiens.
XX WO2004072265-A2.
XX 26-AUG-2004.
XX 11-FEB-2004; 2004WO-US004118.
XX 11-FEB-2003; 2003US-0446133P.
PR 03-APR-2003; 2003US-0459782P.
PR 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
PA (BURC/) BURCZYNSKI M.
PA (TWIN/) TWINE N.
PA (DOR/) DORNER A. J.
PA (TREP/) TREPICCHIO W. L.
XX Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
PI WPI; 2004-642301/62.
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX Disclosure; SEQ ID NO 332; 136pp; English.
XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at
CC least one gene, where the at least one gene is differentially expressed
CC in peripheral blood mononuclear cells (PBMCs) of patients who have a non-
CC blood disease and are subjected to a drug therapy as compared to PBMCs
CC isolated from the patient before the drug therapy, and where the patient
CC has the non-blood disease and is being treated by the drug therapy. The

CC method, kit, and nucleic acid array are useful for monitoring drug
CC activities in vivo. The drug is especially CCI-779, an ester analogue of
CC the immunosuppressant rapamycin which is a potent inhibitor of the
CC mammalian target of rapamycin (mTOR). This sequence represents a gene
CC expressed in PBMC altered by the drug therapy. (Note: this sequence does
CC no form part of the printed specification but was obtained in electronic
CC format from WIPO at ftp.wipo.int/pub/published_pct_sequences/).

SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;
Query Match 100.0%; Score 662; DB 13; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCCGACGAGACCCCTCTCTGACGCCAGCCGCCGCCACCATGSCCAGATT 60
Db 1 ACCGCCGACGAGACCCCTCTCTGACGCCAGCCGCCGCCACCATGSCCAGATT 60
QY 61 CAGCAGCTGGAAGGAGATGGCGCTGCTGACGACGACCAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTGCTGACGACGACCAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGAGTGGGAATAGCTTTTGGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
Db 121 GAGCTAGAGTGGGAATAGCTTTTGGAAAAATGGCGCAATGGCCAGCAGATTGTATC 180
QY 181 ATCACTTGTGATGTTAAAAAGCTCACCATAAAAACTGAGAGCAGCTTTGAAAAACAACACAG 240
Db 181 ATCACTTGTGATGTTAAAAAGCTCACCATAAAAACTGAGAGCAGCTTTGAAAAACAACACAG 240
QY 241 TTTTCTTGTACCTCGGAGAGAGGTTTGAAGAAACACAGCTGTGTCAGAGAAAACTCAG 300
Db 241 TTTTCTTGTACCTCGGAGAGAGGTTTGAAGAAACACAGCTGTGTCAGAGAAAACTCAG 300
QY 301 ACTGTCGCAACTTTACAGATGTCATTTGGTTTCAGCATCAGGAGTGGATGGAGGAA 360
Db 301 ACTGTCGCAACTTTACAGATGTCATTTGGTTTCAGCATCAGGAGTGGATGGAGGAA 360
QY 361 AGCAATAACAAGAAAAATTCAGAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
Db 361 AGCAATAACAAGAAAAATTCAGAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCTTGGACAG 480
Db 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCTTGGACAG 480
QY 481 GAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTT 540
Db 481 GAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTT 540
QY 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCATATAACATTTTACATGAGCTAT 600
Db 541 CTTTCTTTTTCATTACTGTGTTCAATTTATCTTTATCATATAACATTTTACATGAGCTAT 600
QY 601 TTCAAAGTGTGTTGATTAAATTAGATCATCCCTTTGGTTAATAATAAATGTTTGTG 660
Db 601 TTCAAAGTGTGTTGATTAAATTAGATCATCCCTTTGGTTAATAATAAATGTTTGTG 660
QY 661 CT 662
Db 661 CT 662
RESULT 13
ID ADP25373 standard; cDNA; 662 BP.
XX ADP25373;
AC ADP25373;
DT 18-NOV-2004 (first entry)
XX PRO polypeptide encoding cDNA SEQ ID NO:487.
DE
XX

with that in a normal patient sample.

Claim 2: SEO ID NO 24: 53pp; English.

The present invention relates to a method (M1) for diagnosing breast cancer in a patient. The method comprises comparing the level of expression of one or more genes or gene products in a biological sample from the patient with that in a normal patient sample, where a difference in the gene expression in the first sample compared to that in the second sample is a diagnostic of the disease. Also claimed are: method (M2) for distinguishing between normal and disease tissues; method (M3) for monitoring the response of a breast cancer patient to treatment with an anti-cancer agent; method (M4) for identifying a compound for treating breast cancer; and an array for distinguishing between normal and disease tissues comprising two or more probes corresponding to genes selected from ADR9895-ADR99121 or comprising two or more polypeptides selected from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR9895-ADR99121 and the gene products are polypeptides selected from ADR99122-ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are useful for distinguishing between normal and disease tissue. M3 is useful for monitoring the response of a breast cancer patient to treatment with an anti-cancer agent. M4 is useful for identifying a compound for treating breast cancer. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

every Match 100.0%; Score 662; DB 13; Length 662;

Best Local Similarity

Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACGGCGAGCGAGACCCCTCTCTGCAAGCCAGCCGCGCAGCCACCCACCATGGCCACAGTT	60
Db	1	ACGGCGAGCGAGACCCCTCTCTGCAAGCCAGCCGCGCAGCCACCCACCATGGCCACAGTT	60
Qy	61	CAGCAGCTGGAAGGAAGATGGCGCTGGTGACAGCAAAAGCCTTGATGAATAACATGAAG	120
Db	61	CAGCAGCTGGAAGGAAGATGGCGCTGGTGACAGCAAAAGCCTTGATGAATAACATGAAG	120
Qy	121	GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCCAGATTTGTATC	180
Db	121	GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGATTTGTATC	180
Qy	181	ATCAGTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTGAAAAACAACAG	240
Db	181	ATCAGTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTGAAAAACAACAG	240
Qy	241	TTTTCTGTACCTGGGAGAGAAGTTTCAAGAAAACACAGCTGATGGCAGAAAAACTCAG	300
Db	241	TTTTCTGTACCTGGGAGAGAAGTTTCAAGAAAACACAGCTGATGGCAGAAAAACTCAG	300
Qy	301	ACTGTCGCAACTTTACAGATGGTGCAATTGGTTCAGCATCAGGAGTGGGATGGGAAGAA	360
Db	301	ACTGTCGCAACTTTACAGATGGTGCAATTGGTTCAGCATCAGGAGTGGGATGGGAAGAA	360
Qy	361	AGCACAATAACAGAAAATTGAAAGATGGGAAATTTAGTGGTGAGTGCTGCATGAACAAT	420
Db	361	AGCACAATAACAGAAAATTGAAAGATGGGAAATTTAGTGGTGAGTGCTGCATGAACAAT	420
Qy	421	GTCACTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCACATCATCACTTTGGACAG	480
Db	421	GTCACTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCACATCATCACTTTGGACAG	480
Qy	481	GAGTTAATTAAGAGAAATGACCAAGCTCAGTTCAATAGACAAATCTCCATACTGTTCTTT	540
Db	481	GAGTTAATTAAGAGAAATGACCAAGCTCAGTTCAATAGACAAATCTCCATACTGTTCTTT	540
Qy	541	CTTTTTTTTTTCAATTACTGTGTCAAATTTCTTTTATCATAAACAATTTTACATCGACGTA	600
Db	541	CTTTTTTTTTTCAATTACTGTGTCAAATTTCTTTTATCATAAACAATTTTACATCGACGTA	600
Qy	601	TTCAAAAGTGTGTGGATTAAATPAGGATCATCCCTTTGGTTAAATAAAATAAATGTGTGTG	660

CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.
XX

SQ Sequence 662 BP; 210 A; 128 C; 150 G; 174 T; 0 U; 0 Other;

Query Match 100.0%; Score 662; DB 13; Length 662;
Best Local Similarity 100.0%; Pred. No. 1.3e-164;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGCAGCAGACCCCTCTCTGACGCGAGCCGCCGCCGCCACCCACCATGGCCACAGTT 60
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
1 ACCGCGCAGCAGACCCCTCTCTGACGCGAGCCGCCGCCGCCACCCACCATGGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAGATGGCGCTGGTGGACAGCAAGGCTTTGATGAATACATGAAG 120
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 CAGCAGCTGGAAGGAGATGGCGCTGGTGGACAGCAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGGCTTTGATGAATACATGAAG 180
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGGCTTTGATGAATACATGAAG 180
QY 181 ATCATTGTGTGATGGTAAAAACCTCACATATAAACTGAGAGCACTTTGAAAAACAACACAG 240
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 ATCATTGTGTGATGGTAAAAACCTCACATATAAACTGAGAGCACTTTGAAAAACAACACAG 240
QY 241 TTTTCTGTACCTGGGAGAGATTGGAAGAAACACAGCTGTGATGGGAGAAAACTCAG 300
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
241 TTTTCTGTACCTGGGAGAGATTGGAAGAAACACAGCTGTGATGGGAGAAAACTCAG 300
QY 301 ACTGCTGCAACTTTACAGATGGTGCATTGGTTTCAGCATCAGAGTGGGATGGAAGGAA 360
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 ACTGCTGCAACTTTACAGATGGTGCATTGGTTTCAGCATCAGAGTGGGATGGAAGGAA 360
QY 361 AGCACAATAACAAGAAAAATTGAAGATGGGAAATTAGTGGGAGTGTGTCATGAACAAT 420
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
361 AGCACAATAACAAGAAAAATTGAAGATGGGAAATTAGTGGGAGTGTGTCATGAACAAT 420
QY 421 GTCACTGTACTCGGATCTATGAAAAAGTAGATAAAAAATTCATCATCATCTTTGGACAG 480
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
421 GTCACTGTACTCGGATCTATGAAAAAGTAGATAAAAAATTCATCATCATCTTTGGACAG 480
QY 481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCATGAGCAAAATCTCCATCTGTTTCTTT 540
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCATGAGCAAAATCTCCATCTGTTTCTTT 540
QY 541 CTTTTTTTTTTCATTACTGTGTTCAATTATCTTTATCATATAAACATTTTACATGCAGCTAT 600
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 CTTTTTTTTTTCATTACTGTGTTCAATTATCTTTATCATATAAACATTTTACATGCAGCTAT 600
QY 601 TTCAAAGTGTGTGATTAATTAGGATCATCCCTTTGGTTAATAATAAATGTGTTGTG 660
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 TTCAAAGTGTGTGATTAATTAGGATCATCCCTTTGGTTAATAATAAATGTGTTGTG 660
QY 661 CT 662
DB | |
661 CT 662

Search completed: July 12, 2005, 18:01:42
Job time : 420.141 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:39 ; Search time 4994.76 Seconds
(without alignments)

9747.064 Million cell updates/sec

Title: US-09-788-074-2

Perfect score: 1279

Sequence: 1 aatggagcaacatctagc.....tctatgagaagtgcattga 1279

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	710.4	55.5	3260	3 BC053042	BC053042 Mus muscu
2	333	26.0	850	6 CB321549	CB321549 AGENCOURT
3	178.4	13.9	1501	6 CB590454	CB590454 AGENCOURT
4	178.2	13.9	658	1 AA435016	AA435016 ve06a06.r
5	176.6	13.8	293	6 BY791922	BY791922 BY791922
6	176.6	13.8	322	5 BY332071	BY332071 BY332071
7	176.6	13.8	326	5 BY331257	BY331257 BY331257
8	176.6	13.8	327	5 BY129217	BY129217 BY129217
9	176.6	13.8	327	5 BY141407	BY141407 BY141407
10	176.6	13.8	339	5 BY107737	BY107737 BY107737
11	176.6	13.8	341	6 BY784951	BY784951 BY784951
12	176.6	13.8	343	5 BY112490	BY112490 BY112490
13	176.6	13.8	347	5 BY112327	BY112327 BY112327
14	176.6	13.8	347	5 BY112363	BY112363 BY112363
15	176.6	13.8	352	5 BY340471	BY340471 BY340471
16	176.6	13.8	357	5 BY090269	BY090269 BY090269
17	176.6	13.8	358	6 BY768634	BY768634 BY768634
18	176.6	13.8	359	5 BY060959	BY060959 BY060959
19	176.6	13.8	363	5 BY088620	BY088620 BY088620
20	176.6	13.8	363	5 BY101057	BY101057 BY101057
21	176.6	13.8	366	5 BY089011	BY089011 BY089011
22	176.6	13.8	367	6 BY770777	BY770777 BY770777
23	176.6	13.8	374	5 BY320215	BY320215 BY320215
24	176.6	13.8	376	5 BY071168	BY071168 BY071168

ALIGNMENTS

RESULT 1
BC053042
LOCUS BC053042.1 GI:31418541
DEFINITION Mus musculus cDNA clone IMAGE:5687273, with apparent retained intron.
ACCESSION BC053042
VERSION BC053042.1
KEYWORDS HTc.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3260)
AUTHORS Klausner, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Stausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richard, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalley, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3260)

Strausberg, R.

Direct Submission

Submitted (02-JUN-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa


```

Db      229  TTTGTAACCTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCGAAGAACTGAGTCTAG 288
      310  CTACAACATCTGTGAAGCGACAGAAGCTTCTAGATTTCAGATTAAATTCGATTAAACA 369
      289  CTACAACATCTGTGAAGCGACAGAAGCTTCTAGATTTCAGATTAAATTCGATTAAACA 348
      370  TGTCTGTACTTACTCCAGGGCTGACTGAAAAAATCTATTATGAGAGTTGATTTGAT 429
      349  TGTCTGTACTTACTCCAGGGCTGACTGAAAAAATCTATTATGAGAGTTGATTTGAT 408
      430  AAATTAGTAAAGTCCAGGACTAAGAAATGAAGA 464
      409  AAATTAGTAAAGTCCAGGACTAAGAAATGAAGA 443

RESULT 3
CB590454      1501 bp      mRNA      linear      EST 03-APR-2003
LOCUS      AGENCOURT_12600409 NIH_MGC_136 Mus musculus cDNA clone
DEFINITION      IMAGE:30290371 5', mRNA sequence.
ACCESSION      CB590454
VERSION
KEYWORDS
SOURCE      CB590454.1 GI:29508310
ORGANISM      Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 1501)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/
TITLES      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs-remail.nih.gov
      Tissue Procurement: Dr. David Rowe
      cDNA Library Preparation: Invitrogen Corp
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LLNL at:
      http://image.llnl.gov
      Plate: NDAM330 row: o column: 20
      High quality sequence stop: 281.
      Location/Qualifiers
        1..1501
          /organism="Mus musculus"
          /mol_type="mRNA"
          /db_xref="taxon:10090"
          /clone="IMAGE:30290371"
          /tissue_type="embryonic limb, maxilla and mandible"
          /lab_host="DH10B (phage-resistant)"
          /clone_lib="NIH MGC 136"
          /note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
      Normalized, full-length enriched library from pool of
      mouse embryonic limb, maxilla and mandible, embryonic day
      17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw
      equivalents from respective days). Cloned directionally,
      oligo-dT primed (5'-GACTAGTCTAGATCGGCGGCCCTT)15-3'.
      Size selected for the >1kb fragments, average insert size
      1.2 kb. Normalization to Cot 7.5. Tissue contributed by
      David Rowe; library constructed by Resgen, Invitrogen
      Corp. Note: this is a NIH MGC Library."

FEATURES
      source
        1..1501
          /organism="Mus musculus"
          /mol_type="mRNA"
          /db_xref="taxon:10090"
          /clone="IMAGE:30290371"
          /tissue_type="embryonic limb, maxilla and mandible"
          /lab_host="DH10B (phage-resistant)"
          /clone_lib="NIH MGC 136"
          /note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
      Normalized, full-length enriched library from pool of
      mouse embryonic limb, maxilla and mandible, embryonic day
      17.5, 18.5 and newborn (mandible (5, 4 and 1 limb and jaw
      equivalents from respective days). Cloned directionally,
      oligo-dT primed (5'-GACTAGTCTAGATCGGCGGCCCTT)15-3'.
      Size selected for the >1kb fragments, average insert size
      1.2 kb. Normalization to Cot 7.5. Tissue contributed by
      David Rowe; library constructed by Resgen, Invitrogen
      Corp. Note: this is a NIH MGC Library."

ORIGIN
      Query Match      13.9%; Score 178.4; DB 6; Length 1501;
      Best Local Similarity 88.2%; Pred. No. 1.7e-39;
      Matches 194; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY      130  AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCATGGCCAGCCAGACTGTATCATTTAC 189
      86  AGGAGTAGGACTGGCTCTTTAGGAAGATGGCTGCATGGCCAGCCAGACTGTATCATTTAC 145
      190  GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAGTGAGAGCAGCTGTGTCTC 249

```

```

Db      146  GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAGTGAGAGCAGCTGTGTCTC 205
      250  TTTGTAACCTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGTCTAG 309
      206  TTTGTAACCTGGGAGAGAGTTTGTATGAACGACAGCTGATGGCAGAAAACTGAGACGGT 265
      310  CTACAACATCTGTGAAGCGACAGAAGCTTCTAGATTTCAGATTTCAC 349
      266  CTGCAACCACTTTCCAAAGAACCGGTGGCCCTTGGGTTCC 305

RESULT 4
AA435016      658 bp      mRNA      linear      EST 29-MAY-1997
LOCUS      ve06a06.r1 Soares mouse NbMH Mus musculus cDNA clone IMAGE:809266
DEFINITION      5', similar to gb:X70100 rna2 M.musculus mall mRNA for keratinocyte
      lipid-binding (MOUSE);, mRNA sequence.
ACCESSION      AA435016
VERSION
KEYWORDS
SOURCE      AA435016.1 GI:2139930
ORGANISM      Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 658)
AUTHORS      Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
      Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
      Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
      Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
      Waterston, R.
TITLES      The WashU-HMI Mouse EST Project
JOURNAL      Unpublished (1996)
COMMENT      Contact: Marra M/Mouse EST Project
      WashU-HMI Mouse EST Project
      Washington University School of Medicine
      4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
      Tel: 314 286 1800
      Fax: 314 286 1810
      Email: mouseest@watson.wustl.edu
      This clone is available royalty-free through LLNL; contact the
      IMAGE Consortium (info@image.llnl.gov) for further information.
      MGI:485610
      Seq primer: -28mi3 rev2 ET from Amersham
      High quality sequence stop: 290.
      Location/Qualifiers
        1..658
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="CS7BL/6J"
          /db_xref="taxon:10090"
          /clone="IMAGE:809266"
          /sex="male"
          /tissue_type="heart"
          /dev_stage="4 weeks"
          /lab_host="DH10B"
          /clone_lib="Soares mouse NbMH"
          /note="Vector: p7T3D-Pac (Pharmacia) with a modified
      polylinker; Site 1: Site 2: Eco RI; 1st strand cDNA
      was primed with a Not I - oligo(dT) primer [5'
      TGTTCACCATCTGAGTGGAGCGCGCCGGAAGTTTTTTTTTTTTTTTTTTT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not
      I and Eco RI sites of the modified p7T3 vector. RNA
      provided by Dr. Minoru Ko, Wayne State Univ. Library
      constructed and normalized by Bento Soares and M.Fatima
      Bonaldo."

FEATURES
      source
        1..658
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="CS7BL/6J"
          /db_xref="taxon:10090"
          /clone="IMAGE:809266"
          /sex="male"
          /tissue_type="heart"
          /dev_stage="4 weeks"
          /lab_host="DH10B"
          /clone_lib="Soares mouse NbMH"
          /note="Vector: p7T3D-Pac (Pharmacia) with a modified
      polylinker; Site 1: Site 2: Eco RI; 1st strand cDNA
      was primed with a Not I - oligo(dT) primer [5'
      TGTTCACCATCTGAGTGGAGCGCGCCGGAAGTTTTTTTTTTTTTTTTTTT
      3']; double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia), digested with Not I and cloned into the Not
      I and Eco RI sites of the modified p7T3 vector. RNA
      provided by Dr. Minoru Ko, Wayne State Univ. Library
      constructed and normalized by Bento Soares and M.Fatima
      Bonaldo."

ORIGIN
      Query Match      13.9%; Score 178.2; DB 1; Length 658;
      Best Local Similarity 95.8%; Pred. No. 1.5e-39;
      Matches 183; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
 |||||
 Db 116 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 175
 |||||
 QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 249
 |||||
 Db 176 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 235
 |||||
 QY 250 TTGTAACTGGGAGAGAGTGTGATGAACAGACAGCTGATGCGAGAAAACCTGAGGTGAG 309
 |||||
 Db 236 TTGTAACTGGGAGAGAGTGTGATGAACAGACAGCTGATGCGAGAAAACCTGAGGTGAG 295
 |||||
 QY 310 CTACACATAC 320
 |||||
 Db 296 CTGCAACTTCC 306
 |||||

RESULT 5
 BY791922
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BY791922 293 bp mRNA linear EST 23-MAR-2004
 Mus musculus cDNA clone L930265F19 5', mRNA sequence.
 BY791922
 BY791922.1 GI:39718561
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 293)

REFERENCE
 AUTHORS
 Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugihara, Y., Saito, R., Osato, N., Nakamura, M., Shibata, Y., Yasunishi, A., Hirozane-Kishikawa, T., Kusakabe, M., Gustincich, S., Beisel, K., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J., and Hayashizaki, Y.

TITLE
 Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 12819125
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.jp/) for further details.

FEATURES
 source
 1..293
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="L930265F19"
 /tissue_type="whole body"
 /dev_stage="17.5 days embryo"
 /clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN
 Query Match 13.8%; Score 176.6; DB 6; Length 293;

Best Local Similarity 95.3%; Pred. No. 3.3e-39;
 Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 189
 |||||
 Db 40 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACTGTATCATTTAC 99
 |||||
 QY 190 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 249
 |||||
 Db 100 GTGTGATGGCAACAACATCACCGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 159
 |||||
 QY 250 TTGTAACTGGGAGAGAGTGTGATGAACAGACAGCTGATGCGAGAAAACCTGAGGTGAG 309
 |||||
 Db 160 TTGTAACTGGGAGAGAGTGTGATGAACAGACAGCTGATGCGAGAAAACCTGAGGTGAG 219
 |||||
 QY 310 CTACACATAC 320
 |||||
 Db 220 CTGCAACTTCC 230
 |||||

RESULT 6
 BY332071
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BY332071 322 bp mRNA linear EST 11-DEC-2002
 Mus musculus cDNA clone L130023G16 5', mRNA sequence.
 BY332071
 BY332071.1 GI:26522992
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 322)

REFERENCE
 AUTHORS
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verdaro, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 1246851

COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACGAGCTGTATGGCAGAAAACCTGAGGTCTAG 309
 Db 251 TTGTAACTGGGAGAGAGTTTGTATGAACGACGAGCTGTATGGCAGAAAACCTGAGGTCTAG 310
 QY 310 CTACAAATAC 320
 Db 311 CTGCACTTCC 321

RESULT 9
 LOCUS BY141407
 DEFINITION BY141407 RIKEN full-length enriched, 17.5 days embryo whole body
 Mus musculus cDNA clone L930146D12 5', mRNA sequence.

ACCESSION BY141407
 VERSION BY141407.1 GI:26276958
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 327)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusica, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tgami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 Location/Qualifiers
 1..327

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="L930146D12"

/tissue_type="whole body"

/dev_stage="17.5 days embryo"

/clone_lib="RIKEN full-length enriched, 17.5 days embryo
 whole body"

ORIGIN

Query Match 13.8%; Score 176.6; DB 5; Length 327;

Best Local Similarity 95.3%; Pred. No. 3.4e-39;

Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGAGTAGGACTGGCTCTTTAGGAAGTGGCTGCCATGGCCAGCCAGACTGATCATTTAC 189

DB 128 AGAGTAGGACTGGCTCTTTAGGAAGTGGCTGCCATGGCCAGCCAGACTGATCATTTAC 187

QY 190 GTGTGATGTCACCAACATCACCTCAAAACCCAGACGACAGTGAACGACGACTGTGTTCTC 249

DB 188 GTGTGATGTCACCAACATCACCTCAAAACCCAGACGACAGTGAACGACGACTGTGTTCTC 247

QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACGAGCTGTATGGCAGAAAACCTGAGGTCTAG 309

DB 248 TTGTAACTGGGAGAGAGTTTGTATGAACGACGAGCTGTATGGCAGAAAACCTGAGGTCTAG 307

QY 310 CTACAAATAC 320

DB 308 CTGCACCTTCC 318

RESULT 10

BY107737

LOCUS BY107737 RIKEN full-length enriched, 15 days embryo whole body Mus

musculus cDNA clone L330014F16 5', mRNA sequence.

DEFINITION BY107737.1 GI:26218354

ACCESSION BY107737

VERSION BY107737

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 339)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusica, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,

Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,

Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,

Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,

Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,

Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,

Numata, K., Okido, T., Pavan, W.J., Perlea, G., Pesole, G.,

Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,

Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,

Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,

Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,

Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,

Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,

Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,

Rogers, J., Birney, E. and Hayashizaki, Y.

Numata, K., Okido, T., Pavan, W. J., Perteau, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

Location/Qualifiers
1. .339
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="I330014F16"
/tissue_type="whole body"
/dev_stage="15 days embryo"
/clone_lib="RIKEN full-length enriched, 15 days embryo whole body"

ORIGIN

Query Match 13.8%; Score 176.6; DB 5; Length 339;
Best Local Similarity 95.3%; Pred. No. 3.5e-39;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCACGACCCAGACTGTATCATTTAC 189
|||||

DB 99 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCACGACCCAGACTGTATCATTTAC 158
|||||

QY 190 GTGTGATGCGAACACATCACGGTCAAAACCGAGACACAGTGAAGACGACTGTGTCTC 249
|||||

Db 159 GTGTGATGCGAACACATCACGGTCAAAACCGAGACACAGTGAAGACGACTGTGTCTC 218
|||||

QY 250 TTGTAACTGGGAGAGAAAGTTTGATGAACACAGCTGATGGCAGAAAACCTGAGTCAAG 309
|||||

Db 219 TTGTAACTGGGAGAGAAAGTTTGATGAACACAGCTGATGGCAGAAAACCTGAGACGGT 278
|||||

QY 310 CTACACATAC 320
|||||

Db 279 CTGACCTTCC 289
|||||

RESULT 11
LOCUS BY784951 341 bp mRNA linear EST 23-MAR-2004
DEFINITION BY784951 RIKEN full-length enriched, 17.5 days embryo whole body
Mus musculus cDNA clone L930190A14 5', mRNA sequence.

ACCESSION BY784951
VERSION BY784951.1 GI:39711590

KEYWORDS EST
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCES 1 (bases 1 to 341)

AUTHORS Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M., Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S., Sugahara, Y., Saito, R., Osato, N., Fukuda, S., Sato, K., Watahiki, A., Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A., Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K., Pavan, W., Aidinis, V., Nakagawara, A., Held, W. A., Iwata, H., Kono, T., Nakauchi, H., Lyons, P., Wells, C., Hume, D. A., Fagioli, M., Hensch, T. K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P., Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.

TITLE Targeting a complex transcriptome: the construction of the mouse full-length cDNA encyclopedia
Genome Res. 13 (6B), 1273-1289 (2003)

JOURNAL 22703353
MEDLINE 12819125
PubMed

COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

FEATURES
Location/Qualifiers
1. .341
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="L930190A14"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo whole body"

ORIGIN
Query Match 13.8%; Score 176.6; DB 6; Length 341;
Best Local Similarity 95.3%; Pred. No. 3.5e-39;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGGAGTAGACTGGCTCTTAGGAAGATGGCTGCCACGACCCAGACTGTATCATTTAC 189
|||||

Db 120 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACAGCTGTATCATTTAC 179
 QY 190 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAAGTGAAGACGACTGTGTCTC 249
 Db 180 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAAGTGAAGACGACTGTGTCTC 239
 QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGCCAGAAAACCTGAGTCCAG 309
 Db 240 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGCCAGAAAACCTGAGACGGT 299
 QY 310 CTCAACATAC 320
 Db 300 CTGCACCTTCC 310

RESULT 12
 BY112490
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BY112490.1 GI:26223107
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 343)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustinch, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagaishina, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

1246681
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

FEATURES
 source

Location/Qualifiers
 1..343
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="L430007B06"
 /tissue_type="whole body"
 /dev_stage="18 days embryo"
 /clone_lib="RIKEN full-length enriched, 18 days embryo
 whole body"

ORIGIN

Query Match 13.8%; Score 176.6; DB 5; Length 343;
 Best Local Similarity 95.3%; Pred. No. 3.5e-39;
 Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACAGCTGTATCATTTAC 189
 Db 127 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAAAGCCAGACAGCTGTATCATTTAC 186
 QY 190 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAAGTGAAGACGACTGTGTCTC 249
 Db 187 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCAAGTGAAGACGACTGTGTCTC 246
 QY 250 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGCCAGAAAACCTGAGTCCAG 309
 Db 247 TTGTAACTGGGAGAGAGTTTGTATGAACGACAGCTGATGCCAGAAAACCTGAGACGGT 306
 QY 310 CTCAACATAC 320
 Db 307 CTGCACCTTCC 317

RESULT 13

BY112327

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.L., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL MEDLINE
Nature 420, 563-573 (2002)

PUBMED
22354683

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source
1. 347
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L430005F10"
/tissue_type="whole body"
/dev_stage="18 days embryo"
/clone_lib="RIKEN full-length enriched, 18 days embryo whole body"

ORIGIN

Query Match 13.8%; Score 176.6; DB 5; Length 347;
Best Local Similarity 95.3%; Pred. No. 3.5e-39;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCATGCCCAAGCCAGACTGTATCATTTAC 189
|||||
Db 131 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCATGCCCAAGCCAGACTGTATCATTTAC 190
|||||
QY 190 GTGTGATGGCAACCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACCTGTGTTCTC 249
|||||
Db 191 GTGTGATGGCAACCAACATCACGGTCAAAACCGAGAGCAGTGAAGACGACCTGTGTTCTC 250
|||||
QY 250 TTGTAACTGGGAGAGAAAGTTTGATGAACACGACAGCTGATGCGCAGAAAACTGAGGTGAG 309
|||||
Db 251 TTGTAACTGGGAGAGAAAGTTTGATGAACACGACAGCTGATGCGCAGAAAACTGAGCGGT 310
|||||
QY 310 CTACAACATATC 320
|||||
Db 311 CTGCACCTTCC 321
|||||

RESULT 14

LOCUS BY112363

DEFINITION BY112363 RIKEN full-length enriched, 18 days embryo whole body Mus

ACCESSION BY112363

VERSION BY112363.1 GI:26222980

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 347)

AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.L., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

BY112363 347 bp mRNA linear EST 07-DEC-2002
BY112363 RIKEN full-length enriched, 18 days embryo whole body Mus
musculus cDNA clone L430005M13 5', mRNA sequence.

BY112363

EST. BY112363.1 GI:26222980

Mus musculus (house mouse)

Mus musculus

1 (bases 1 to 347)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, D.J., Reid, J., Ring, B.L., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sakai, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES source
Location/Qualifiers
1. 347
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L430005M13"
/tissue_type="whole body"
/dev_stage="18 days embryo"
/clone_lib="RIKEN full-length enriched, 18 days embryo whole body"

ORIGIN
Query Match 13.8%; Score 176.6; DB 5; Length 347;
Best Local Similarity 95.3%; Pred. No. 3.5e-39;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 130 AGGAGTAGGACCTGGCTCTTAGGAAGATGGCTGGCCAGGACGACGATGATCATTTAC 189
DB 131 AGGAGTAGGACCTGGCTCTTAGGAGATGGCTGGCCAGGACGACGATGATCATTTAC 190
QY 190 GTGTGATGGCAACATCAACGTCACCGAGACGACGATGAGACGACGATGTTCTC 249
DB 191 GTGTGATGGCAACATCAACGTCACCGAGACGACGATGAGACGACGATGTTCTC 250
QY 250 TTGTAACTGGGAGAGAGTTGATGAACGACGACGATGATGCGAGAAAACTGAGTCAG 309
DB 251 TTGTAACTGGGAGAGAGTTGATGAACGACGACGATGATGCGAGAAAACTGAGTCAGT 310
QY 310 CTACACATAC 320
DB 311 CTGCACCTTCC 321

RESULT 15
LOCUS BY340471 352 bp mRNA linear EST 12-DBC-2002
DEFINITION BY340471 RIKEN full-length enriched, whole joints Mus musculus cDNA clone L230028H20 5', mRNA sequence.
ACCESSION BY340471
VERSION BY340471.1 GI:26569959
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 352)
AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuoka, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brucic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltava, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sadelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomica, M., Verdano, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Waki, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Vassilis Aidinis (Biomedical Sciences Research Center 'Al. Fleming' Institute of Immunology 14-16 Al. Fleming street 16672 Vari, Greece) whose assistance we gratefully acknowledge.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES source
Location/Qualifiers
1. 352
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="L230028H20"
/tissue_type="whole joints"
/clone_lib="RIKEN full-length enriched, whole joints"

ORIGIN

Query Match 13.8%; Score 176.6; DB 5; Length 352;
Best Local Similarity 95.3%; Pred. No. 3.5e-39;
Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 130 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTAC 189
|||
Db 127 AGGAGTAGGACTGGCTCTTAGGAAGATGGCTGCCATGGCCAGCCAGACTGTATCATTAC 186
|||

Qy 190 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 249
|||
Db 187 GTGTGATGGCAACAACATCACGGTCAAAACCGAGAGCACAGTGAAGACGACTGTGTTCTC 246
|||

Qy 250 TTGTAACTGGGAGAGAGTTTGTGAAACGACAGCTGATGGCAGAAAACTGAGGTCAG 309
|||
Db 247 TTGTAACTGGGAGAGAGTTTGTGAAACGACAGCTGATGGCAGAAAACTGAGACGGT 306
|||

Qy 310 CTACAACATAC 320
|||
Db 307 CTGCACCTTCC 317
|||

Search completed: July 12, 2005, 21:11.48
Job time : 4999.76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:38 ; Search time 1167.11 Seconds
(without alignments)
3561.142 Million cell updates/sec

Title: US-09-788-074-4
Perfect score: 662
Sequence: 1 accgcgcagcagaccctc.....taataaatgtttgtgtgt 662

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6330945 seqs, 3139162390 residues

Total number of hits satisfying chosen parameters: 12661890

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA.*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
 - 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
 - 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
 - 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
 - 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
 - 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
 - 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
 - 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	662	100.0	662	9	US-09-788-074-4
2	662	100.0	662	15	US-10-171-581-95
3	662	100.0	662	15	US-10-205-823-102
4	662	100.0	662	17	US-10-170-385-210
5	662	100.0	662	17	US-10-172-118-608
6	662	100.0	662	18	US-10-342-887-608
7	662	100.0	662	19	US-10-717-597-57

8	662	100.0	662	19	US-10-775-169-332	Sequence 332, App
9	662	100.0	662	19	US-10-788-792-24	Sequence 24, Appl
10	660	99.7	1071	10	US-09-971-392-113	Sequence 113, App
11	660	99.7	1071	10	US-09-971-429B-52	Sequence 52, Appl
12	649	98.0	1072	16	US-10-240-965-143	Sequence 143, App
13	590	89.1	615	9	US-09-920-455-23	Sequence 23, Appl
14	583	88.1	606	9	US-09-920-455-147	Sequence 147, App
15	535.6	80.9	634	10	US-09-981-151A-17	Sequence 17, Appl
16	517	78.1	555	9	US-09-920-455-56	Sequence 56, Appl
17	512	77.3	519	9	US-09-920-455-189	Sequence 189, App
18	500	75.5	566	9	US-09-920-455-69	Sequence 69, Appl
19	476.4	72.0	494	10	US-09-918-995-23104	Sequence 23104, A
20	474.2	71.6	588	10	US-09-918-995-13963	Sequence 14963, A
21	470	71.0	471	17	US-10-242-535A-24416	Sequence 24416, A
22	460.8	69.6	491	11	US-09-972-211-9	Sequence 9, Appl1
23	460.8	69.6	491	18	US-10-096-625-9	Sequence 9, Appl1
24	460.8	69.6	491	18	US-09-920-455-175	Sequence 175, App
25	451	68.1	461	9	US-09-804-014A-44	Sequence 44, Appl
26	444	67.1	444	10	US-10-029-386-4013	Sequence 4013, Ap
27	439.4	66.4	501	16	US-10-242-535A-7126	Sequence 7126, Ap
28	439	66.3	439	17	US-10-085-783A-7126	Sequence 7126, Ap
29	439	66.3	439	18	US-09-918-995-25588	Sequence 25588, A
30	429.8	64.9	479	10	US-10-242-535A-20520	Sequence 20520, A
31	407.6	61.6	440	17	US-10-085-783A-20520	Sequence 20520, A
32	407.6	61.6	440	18	US-10-161-927-19	Sequence 19, Appl
33	406	61.3	435	17	US-10-027-632-289654	Sequence 289654,
34	406	61.3	435	17	US-10-027-632-289655	Sequence 289655,
35	391.8	59.2	620	13	US-10-027-632-289656	Sequence 289656,
36	391.8	59.2	620	13	US-10-027-632-289657	Sequence 289657,
37	391.8	59.2	620	13	US-10-027-632-289657	Sequence 289657,
38	391.8	59.2	620	17	US-10-027-632-289657	Sequence 289657,
39	391.8	59.2	620	17	US-10-027-632-289657	Sequence 289657,
40	391.8	59.2	620	17	US-10-027-632-289657	Sequence 289657,
41	372.6	56.3	602	13	US-10-027-632-289657	Sequence 289657,
42	372.6	56.3	602	17	US-10-027-632-289657	Sequence 289657,
43	370.6	56.0	835	13	US-10-027-632-165740	Sequence 165740,
44	370.6	56.0	835	17	US-10-027-632-165740	Sequence 165740,
45	367	55.4	413	10	US-09-804-014A-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1

US-09-788-074-4
; Sequence 4, Application US/09788074
; Patent No. US20010044110A1
; GENERAL INFORMATION:
; APPLICANT: Hotamisligil, Gokhan S.
; TITLE OF INVENTION: INHIBITION OF MAL1
; FILE REFERENCE: 21509-044
; CURRENT APPLICATION NUMBER: US/09/788,074
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,106
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-788-074-4

Query Match	100.0%	Score 662;	DB 9;	Length 662;
Best Local Similarity	100.0%	Pred. No. 2.3e-162;		
Matches 662;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACCGCGCAGCAGACCCCTCTCTGCACGCAGCCGCCGCCGCCACCCACCCAGTT	60	
Db	1	ACCGCGCAGCAGACCCCTCTCTGCACGCAGCCGCCGCCGCCACCCACCCAGTT	60	
QY	61	CAGCAGCTGGAAGGAGAGTGGCGCTGTGTGGACAGCAAGGCTTTGATGATCATGAAG	120	

Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
Qy |||||
Db 121 GAGCTAGAGTGGGATAGCTTTGCGAATAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Qy |||||
Db 121 GAGCTAGAGTGGGATAGCTTTGCGAATAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Qy |||||
Db 181 ATCACTTGTGATGTTAAACCTCACCATATAAACTGAGAGCACTTTGAAACCAACACAG 240
Qy |||||
Db 181 ATCACTTGTGATGTTAAACCTCACCATATAAACTGAGAGCACTTTGAAACCAACACAG 240
Qy |||||
Db 241 TTTTCTGTACCTCGGAGAGAGTTTGAAGAAACCAAGCTGTGAGCAATGGGATGGGAGGAA 300
Qy |||||
Db 241 TTTTCTGTACCTCGGAGAGAGTTTGAAGAAACCAAGCTGTGAGCAATGGGATGGGAGGAA 300
Qy |||||
Db 301 ACTGTCTGCACTTTACAGATGGTGCATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
Qy |||||
Db 301 ACTGTCTGCACTTTACAGATGGTGCATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
Qy |||||
Db 421 GTCACTGTACTCGGATCTATGAAAGATGAGAAATTTCCATCATCATCTTTGGACAG 480
Qy |||||
Db 421 GTCACTGTACTCGGATCTATGAAAGATGAGAAATTTCCATCATCATCTTTGGACAG 480
Qy |||||
Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATCATCTTTGGTCTTTT 540
Qy |||||
Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATCATCTTTGGTCTTTT 540
Qy |||||
Db 541 CTTTCTTTTCTCACTACTGTGTCAATATCTTTATCATATAACATTTTACATGAGCTAT 600
Qy |||||
Db 541 CTTTCTTTTCTCACTACTGTGTCAATATCTTTATCATATAACATTTTACATGAGCTAT 600
Qy |||||
Db 601 TTCAAAGTGTGTGGATTAATAGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
Qy |||||
Db 601 TTCAAAGTGTGTGGATTAATAGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
Qy 661 CT 662
Db 661 CT 662

RESULT 2

US-10-171-581-95
; Sequence 95, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 95
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: M94856
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-95

Query Match 100.0%; Score 662; DB 15; Length 662;
Best Local Similarity 100.0%; Pred. No. 2,3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCACCCACCATGCCACAGTT 60
|||||

Db 1 ACCGCGGACGAGACCCCTCTCTGACGCGCCGCGCACCCACCATGCCACAGTT 60
Qy |||||
Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
Qy |||||
Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
Qy |||||
Db 121 GAGCTAGAGTGGGATAGCTTTGCGAATAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Qy |||||
Db 121 GAGCTAGAGTGGGATAGCTTTGCGAATAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Qy |||||
Db 181 ATCACTTGTGATGTTAAACCTCACCATATAAACTGAGAGCACTTTGAAACCAACACAG 240
Qy |||||
Db 181 ATCACTTGTGATGTTAAACCTCACCATATAAACTGAGAGCACTTTGAAACCAACACAG 240
Qy |||||
Db 241 TTTTCTGTACCTCGGAGAGAGTTTGAAGAAACCAAGCTGTGAGCAATGGGATGGGAGGAA 300
Qy |||||
Db 241 TTTTCTGTACCTCGGAGAGAGTTTGAAGAAACCAAGCTGTGAGCAATGGGATGGGAGGAA 300
Qy |||||
Db 301 ACTGTCTGCACTTTTACAGATGGTGCATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 360
Qy |||||
Db 301 ACTGTCTGCACTTTTACAGATGGTGCATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 360
Qy |||||
Db 361 AGCAATATAACAGAAATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
Qy |||||
Db 361 AGCAATATAACAGAAATTTGAAAGATGGGAAATTTAGTGGAGTGTGTATGAACAAT 420
Qy |||||
Db 421 GTCACTGTACTCGGATCTATGAAAGATGAGAAATTTCCATCATCATCTTTGGACAG 480
Qy |||||
Db 421 GTCACTGTACTCGGATCTATGAAAGATGAGAAATTTCCATCATCATCTTTGGACAG 480
Qy |||||
Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATCATCTTTGGTCTTTT 540
Qy |||||
Db 481 GAGTTAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAATCTCCATCATCTTTGGTCTTTT 540
Qy |||||
Db 541 CTTTCTTTTCTCACTACTGTGTCAATATCTTTATCATATAACATTTTACATGAGCTAT 600
Qy |||||
Db 541 CTTTCTTTTCTCACTACTGTGTCAATATCTTTATCATATAACATTTTACATGAGCTAT 600
Qy |||||
Db 601 TTCAAAGTGTGTGGATTAATAGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
Qy |||||
Db 601 TTCAAAGTGTGTGGATTAATAGATCATCCCTTTGGTTAATAAATGTTGTTGTG 660
Qy 661 CT 662
Db 661 CT 662

RESULT 3

US-10-205-823-102
; Sequence 102, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Wonsay, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22

```
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 102
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-823-102

Query Match      100.0%; Score 662; DB 15; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGCGCAGACCCCTCTCTGACGCGCCGCGCCGCGCCACCCACCATGGCCACAGTT 60
DB 1 ACGCGGCGCAGACCCCTCTCTGACGCGCCGCGCCGCGCCACCCACCATGGCCACAGTT 60

QY 61 CAGCAGCTCGAAGGAAGATGGCGCTCTGCGAGCAGCAAAAGGCTTTGATGAATACATGAAG 120
DB 61 CAGCAGCTCGAAGGAAGATGGCGCTCTGCGAGCAGCAAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
DB 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180

QY 181 ATCACTTGTGATGTTAAACCTCACCATAAAACCTGAGAGCATTGTAACCAACACAG 240
DB 181 ATCACTTGTGATGTTAAACCTCACCATAAAACCTGAGAGCATTGTAACCAACACAG 240

QY 241 TTTTCTGTACCTGGGAGAGATGTTGAAAGAAACCAACAGCTGATGGCGAGAAAACCTCAG 300
DB 241 TTTTCTGTACCTGGGAGAGATGTTGAAAGAAACCAACAGCTGATGGCGAGAAAACCTCAG 300

QY 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGCGAAAAATGGCGCAATGGCCAAAGCTGATC 360
DB 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGCGAAAAATGGCGCAATGGCCAAAGCTGATC 360

QY 361 AGCACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAAT 420
DB 361 AGCACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAAT 420

QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAATTCATCATCTTGGACAG 480
DB 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAATTCATCATCTTGGACAG 480

QY 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATTAATAAATTCATCATCTTGGACAG 540
DB 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATTAATAAATTCATCATCTTGGACAG 540

QY 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATTAATTAATGTTTGG 660
DB 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATTAATTAATGTTTGG 660

QY 661 CT 662
DB 661 CT 662
```

RESULT 4

```
US-10-170-385-210
; Sequence 210, Application US/10170385
; Publication No. US2003020372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
```

```
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 53268200100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 210
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-170-385-210

Query Match      100.0%; Score 662; DB 17; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGCGGCGCAGACCCCTCTCTGACGCGCCGCGCCGCGCCACCCACCATGGCCACAGTT 60
DB 1 ACGCGGCGCAGACCCCTCTCTGACGCGCCGCGCCGCGCCACCCACCATGGCCACAGTT 60

QY 61 CAGCAGCTGGAAGGAAGATGGCGCTCTGCGAGCAGCAAAAGGCTTTGATGAATACATGAAG 120
DB 61 CAGCAGCTGGAAGGAAGATGGCGCTCTGCGAGCAGCAAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCTGATGTATC 180
DB 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAAAGCTGATGTATC 180

QY 181 ATCACTTGTGATGTTAAACCTCACCATAAAACCTGAGAGCATTGTAACCAACACAG 240
DB 181 ATCACTTGTGATGTTAAACCTCACCATAAAACCTGAGAGCATTGTAACCAACACAG 240

QY 241 TTTTCTGTACCTGGGAGAGATGTTGAAAGAAACCAACAGCTGATGGCGAGAAAACCTCAG 300
DB 241 TTTTCTGTACCTGGGAGAGATGTTGAAAGAAACCAACAGCTGATGGCGAGAAAACCTCAG 300

QY 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGCGAAAAATGGCGCAATGGCCAAAGCTGATC 360
DB 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGCGAAAAATGGCGCAATGGCCAAAGCTGATC 360

QY 361 AGCACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAAT 420
DB 361 AGCACAATAACAAGAAATTTGAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAAT 420

QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAATTCATCATCTTGGACAG 480
DB 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAATTCATCATCTTGGACAG 480

QY 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATTAATAAATTCATCATCTTGGACAG 540
DB 481 GAGTTAAATTAAGAGAAATGACCAAGCTCAGTTCAATTAATAAATTCATCATCTTGGACAG 540

QY 541 CTTTCTTTTTCATTTACTGTGTTCATTTATCTTTATCATATAACATTTTACATGCAGCTAT 600
DB 541 CTTTCTTTTTCATTTACTGTGTTCATTTATCTTTATCATATAACATTTTACATGCAGCTAT 600

QY 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATTAATTAATGTTTGG 660
DB 601 TTCAAAGTGTGTGGATTAATAGGATCATCCCTTTGGTTAATAAATTAATTAATGTTTGG 660
```

```
QY 661 CT 662
Db 661 CT 662

RESULT 5
US-10-172-118-608
; Sequence 608, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 608
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 001444
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-608

Query Match 100.0%; Score 662; DB 17; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGACGAGACCCCTCTCTGACGCGCAGCCCGCGGACCCGACCCACCATGGCCACAGTT 60
Db 1 ACCGCGGACGAGACCCCTCTCTGACGCGCAGCCCGCGGACCCGACCCACCATGGCCACAGTT 60

QY 61 CAGCAGCTGGAAGGAGATGGCGCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTTGGGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTTGGGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180

QY 181 ATCACTTTGTGATGTTAAACCTCACCATAAAAACTGAGAGACATTTGAAAAACAACAG 240
Db 181 ATCACTTTGTGATGTTAAACCTCACCATAAAAACTGAGAGACATTTGAAAAACAACAG 240

QY 241 TTTTCTGTACCTCGGAGAGATTTGAAAGAACCCACAGCTGTATGGCGAGAAAACTCAG 300
Db 241 TTTTCTGTACCTCGGAGAGATTTGAAAGAACCCACAGCTGTATGGCGAGAAAACTCAG 300

QY 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGAA 360
Db 301 ACTGTCTGCAACTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGAA 360

QY 361 AGCACAATAACAAGAAATTTGAAAGTAGAATAAAAAATTTAGTGGTGGTGTCTATGAACAT 420
Db 361 AGCACAATAACAAGAAATTTGAAAGTAGAATAAAAAATTTAGTGGTGGTGTCTATGAACAT 420

QY 421 GTCACTGTACTCGGATCTATGAAAGTAGAATAAAAAATTTCCATCATCACTTTGGACAG 480
Db 421 GTCACTGTACTCGGATCTATGAAAGTAGAATAAAAAATTTCCATCATCACTTTGGACAG 480

QY 481 GAGTTAATTAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTT 540
Db 481 GAGTTAATTAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTT 540

US-10-342-887-608
; Sequence 608, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 608
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-608

Query Match 100.0%; Score 662; DB 18; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGCGGACGAGACCCCTCTCTGACGCGCAGCCCGCGGACCCGACCCACCATGGCCACAGTT 60
Db 1 ACCGCGGACGAGACCCCTCTCTGACGCGCAGCCCGCGGACCCGACCCACCATGGCCACAGTT 60

QY 61 CAGCAGCTGGAAGGAGATGGCGCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120

QY 121 GAGCTAGGAGTGGGAATAGCTTTTGGGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTTGGGAAAAATGGCGCAATGGCCAAAGCCAGATTGTATC 180

QY 181 ATCACTTTGTGATGTTAAACCTCACCATAAAAACTGAGAGACATTTGAAAAACAACAG 240
Db 181 ATCACTTTGTGATGTTAAACCTCACCATAAAAACTGAGAGACATTTGAAAAACAACAG 240

QY 241 TTTTCTGTACCTCGGAGAGATTTGAAAGAACCCACAGCTGTATGGCGAGAAAACTCAG 300
Db 241 TTTTCTGTACCTCGGAGAGATTTGAAAGAACCCACAGCTGTATGGCGAGAAAACTCAG 300

QY 301 ACTGTCTGCAACTTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGAA 360
Db 301 ACTGTCTGCAACTTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGAA 360

QY 361 AGCACAATAACAAGAAATTTGAAAGTAGAATAAAAAATTTAGTGGTGGAGTGTCTATGAACAT 420
```

Db 361 AGACATTAACAGAGAAATGAAAGATGGAAATAGTGGTGGAGTGTGCATGAACAAT 420
QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAAAATCCATCATCACTTTGGACAG 480
Db 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAAAATCCATCATCACTTTGGACAG 480
QY 481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTTTTCTTT 540
Db 481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTTTTCTTT 540
QY 541 CTTTCTTTTTCATTAATCTGTTCAATTAATCTTTATCATAAACAATTTTACATGCAGCTAT 600
Db 541 CTTTCTTTTTCATTAATCTGTTCAATTAATCTTTATCATAAACAATTTTACATGCAGCTAT 600
QY 601 TTCAAAGTGTGTGGATTAATAGGATCAATCCCTTTGGTTAATAAATAATGTTTGTG 660
Db 601 TTCAAAGTGTGTGGATTAATAGGATCAATCCCTTTGGTTAATAAATAATGTTTGTG 660
QY 661 CT 662
Db 661 CT 662

RESULT 7

US-10-717-597-57
; Sequence 57, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dörner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 57
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-57

Query Match 100.0%; Score 662; DB 19; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCACCCACCATGCGCCACAGTT 60
Db 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCACCCACCATGCGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCCAGATTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCCAGATTGTATC 180
QY 181 ATCACTTGTGATGGTAAAAACCTCACCATATAAAAACTGAGAGCACTTTTGAACCAACACAG 240
Db 181 ATCACTTGTGATGGTAAAAACCTCACCATATAAAAACTGAGAGCACTTTTGAACCAACACAG 240
QY 241 TTTTCTGTACCTCGGAGAGAAAGTTTGAAGAAACCAACAGCTGATGGCGAGAAAAAATCTCAG 300

Db 241 TTTTCTGTACCTCGGAGAGAAAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAATCTCAG 300
QY 301 ACTGCTGTCAAACTTTTACAGATGGTGCATTTGGTTTACAGCATCAGGAGTGGATGGGAAGGAA 360
Db 301 ACTGCTGTCAAACTTTTACAGATGGTGCATTTGGTTTACAGCATCAGGAGTGGATGGGAAGGAA 360
QY 361 AGCAACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTATGAACAAT 420
Db 361 AGCAACAATAACAAGAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCTATGAACAAT 420
QY 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAAAATCCATCATCACTTTTGGACAG 480
Db 421 GTCACTGTACTCGGATCTATGAAAAGTAGAATAAAAAATCCATCATCACTTTTGGACAG 480
QY 481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTTTTCTTT 540
Db 481 GAGTTAAATTAAGAGATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCACTTTTCTTT 540
QY 541 CTTTCTTTTTCATTAATCTGTTCAATTAATCTTTATCATAAACAATTTTACATGCAGCTAT 600
Db 541 CTTTCTTTTTCATTAATCTGTTCAATTAATCTTTATCATAAACAATTTTACATGCAGCTAT 600
QY 601 TTCAAAGTGTGTGGATTAATAGGATCAATCCCTTTGGTTAATAAATAATGTTTGTG 660
Db 601 TTCAAAGTGTGTGGATTAATAGGATCAATCCCTTTGGTTAATAAATAATGTTTGTG 660
QY 661 CT 662
Db 661 CT 662

RESULT 8

US-10-775-169-332
; Sequence 332, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775,169
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 332
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-332

Query Match 100.0%; Score 662; DB 19; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCACCCACCATGCGCCACAGTT 60
Db 1 ACCGCGGACGACAGCCCTCTCTGACGCGCAGCCGCGCACCCACCATGCGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCCAGATTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCAGCCAGATTGTATC 180
QY 181 ATCACTTGTGATGGTAAAAACCTCACCATATAAAAACTGAGAGCACTTTTGAACCAACACAG 240
Db 181 ATCACTTGTGATGGTAAAAACCTCACCATATAAAAACTGAGAGCACTTTTGAACCAACACAG 240

```
QY 241 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAAGTCTCAG 300
Db 241 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAAGTCTCAG 300
QY 301 ACTGCTGCAACTTTTACAGATGGTGCATTGGTTTACAGCATCAGGAGTGGGATGGGAGGAA 360
Db 301 ACTGCTGCAACTTTTACAGATGGTGCATTGGTTTACAGCATCAGGAGTGGGATGGGAGGAA 360
QY 361 AGCAACAATAACAAGAAAAATTGAAGATGGGAAATTTAGTGGAGTGTGTTCATGAACAAT 420
Db 361 AGCAACAATAACAAGAAAAATTGAAGATGGGAAATTTAGTGGAGTGTGTTCATGAACAAT 420
QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
Db 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
QY 481 GAGTTAAATTAAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTTCTTT 540
Db 481 GAGTTAAATTAAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTTCTTT 540
QY 541 CTTTCTTTTTCATTTACTGTGTCAATTATCTTTATCATATAAACAATTTTACATGCAGCTAT 600
Db 541 CTTTCTTTTTCATTTACTGTGTCAATTATCTTTATCATATAAACAATTTTACATGCAGCTAT 600
QY 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTG 660
Db 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTG 660
QY 661 CT 662
Db 661 CT 662
```

RESULT 9

```
US-10-788-792-24
; Sequence 24, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-788-792-24
```

```
Query Match 100.0%; Score 662; DB 19; Length 662;
Best Local Similarity 100.0%; Pred. No. 2.3e-162;
Matches 662; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ACCGCGGACGACAGCCCTCTCTGCAGCGCCGCCGCCGCCACCATGGCCACAGTT 60
Db 1 ACCGCGGACGACAGCCCTCTCTGCAGCGCCGCCGCCGCCACCATGGCCACAGTT 60
QY 61 CAGCAGCTGGAAGGAGATGGCGCTCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120
Db 61 CAGCAGCTGGAAGGAGATGGCGCTCTGTGGACGACAAAGGCTTTGATGAATACATGAAG 120
QY 121 GAGCTAGGAGTGGGAATAGCTTTGGAAAAAATGGGCGCAATGGCCAGGCAGATTGTATC 180
Db 121 GAGCTAGGAGTGGGAATAGCTTTGGAAAAAATGGGCGCAATGGCCAGGCAGATTGTATC 180
QY 181 ATCAGTTGTGATGGTAAAAAAGCTTCAACATAAAAACTGAGAGCAGCTTTTGAACAAACACAG 240
```

```
Db 181 ATCAGTTGTGATGGTAAAAAAGCTTCAACATAAAAACTGAGAGCAGCTTTTGAACAAACACAG 240
QY 241 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAAGTCTCAG 300
Db 241 TTTCTTGTACCTGGGAGAGAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAAGTCTCAG 300
QY 301 ACTGCTGCAACTTTTACAGATGGTGCATTGGTTTACAGCATCAGGAGTGGGATGGGAGGAA 360
Db 301 ACTGCTGCAACTTTTACAGATGGTGCATTGGTTTACAGCATCAGGAGTGGGATGGGAGGAA 360
QY 361 AGCAACAATAACAAGAAAAATTGAAGATGGGAAATTTAGTGGAGTGTGTTCATGAACAAT 420
Db 361 AGCAACAATAACAAGAAAAATTGAAGATGGGAAATTTAGTGGAGTGTGTTCATGAACAAT 420
QY 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
Db 421 GTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAG 480
QY 481 GAGTTAAATTAAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTTCTTT 540
Db 481 GAGTTAAATTAAAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATACCTGTTTCTTT 540
QY 541 CTTTCTTTTTCATTTACTGTGTCAATTATCTTTATCATATAAACAATTTTACATGCAGCTAT 600
Db 541 CTTTCTTTTTCATTTACTGTGTCAATTATCTTTATCATATAAACAATTTTACATGCAGCTAT 600
QY 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTG 660
Db 601 TTCAAAGTGTGTGGATTAATTAGGATCATCCCTTTGGTTAATAAATAAATGTGTTTGTG 660
QY 661 CT 662
Db 661 CT 662
```

RESULT 10

```
US-09-971-392-113
; Sequence 113, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 113
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Template ID: 995880.12
US-09-971-392-113
```

```
Query Match 99.7%; Score 660; DB 10; Length 1071;
Best Local Similarity 100.0%; Pred. No. 9.6e-162;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3 CGCCGACGACAGACCCCTCTCTGCAGCGCCGCCGCCGCCACCATGGCCACAGTTCA 62
Db 42 CGCCGACGACAGACCCCTCTCTGCAGCGCCGCCGCCGCCACCATGGCCACAGTTCA 101
QY 63 GCAGCTGGAAGAGAGATGGCGCTCTGTGGACGACGAAAGGCTTTGATGATACATGAGCA 122
Db 102 GCAGCTGGAAGAGAGATGGCGCTCTGTGGACGACGAAAGGCTTTGATGATACATGAGCA 161
```

123	Qy	GCTAGGAGTGGGAATAGCTTTGCGAAAAAATGGCGCGCAATGGCCAAAGCCAGATGTTGATCAT	182
162	Db	GCTAGGAGTGGGAATAGCTTTGCGAAAAAATGGCGCGCAATGGCCAAAGCCAGATGTTGATCAT	221
183	Qy	CAC TTGTGATGGTAAAAACCTCACCATAAAAACTCAGAGCAGCTTTTGAAAAACACACAGTT	242
222	Db	CAC TTGTGATGGTAAAAACCTCACCATAAAAACTCAGAGCAGCTTTTGAAAAACACACAGTT	281
243	Qy	TTCTTTGTACCTGGGAGAGAGAAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAATCTCAGAC	302
282	Db	TTCTTTGTACCTGGGAGAGAGAAGTTTGAAGAAACCAACAGCTGATGGCAGAAAAAATCTCAGAC	341
303	Qy	TGTC TGCAACTTTACAGATGGTGCAATGGTTCAGCATCAGGAGTGGGATGGGAAGGAAAG	362
342	Db	TGTC TGCAACTTTACAGATGGTGCAATGGTTCAGCATCAGGAGTGGGATGGGAAGGAAAG	401
363	Qy	CACAATAACAAGAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTGCATGAACAATGT	422
402	Db	CACAATAACAAGAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTGCATGAACAATGT	461
423	Qy	CACCTGTACTCGGACTATGAAAAAGTAGATAAAAAATTCACATCATCACTTTGGACAGGA	482
462	Db	CACCTGTACTCGGACTATGAAAAAGTAGATAAAAAATTCACATCATCACTTTGGACAGGA	521
483	Qy	GTTAAATTAAGAGATGACCAAGCTCAGTTCAATAGAGCAAACTCTCCATCTGTTTCTTTCT	542
522	Db	GTTAAATTAAGAGATGACCAAGCTCAGTTCAATAGAGCAAACTCTCCATCTGTTTCTTTCT	581
543	Qy	TTTTTTTTTTCATTA CTGTTGTTTCAATTA TCTTTATCATAAACATTTTACATGCAGCTATTT	602
582	Db	TTTTTTTTTTCATTA CTGTTGTTTCAATTA TCTTTATCATAAACATTTTACATGCAGCTATTT	641
603	Qy	CAAAAGTGTGGATTAATAGGATCATCCCTTTGGTTTAATAAATAAATGTTGTGTGCT	662
642	Db	CAAAAGTGTGGATTAATAGGATCATCCCTTTGGTTTAATAAATAAATGTTGTGTGCT	701

RESULT 11

```

RES001 11
US-09-971-429B-52
; Sequence 52, Application US/09971429B
; Publication No. US20030175704A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K. W.
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
; FILE REFERENCE: PA-0040 US
; CURRENT APPLICATION NUMBER: US/09/971,429B
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/239,024
; PRIOR FILING DATE: 2000-04-10
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 52
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030175704A1 995880.12
US-09-971-429B-52

```

	Query Match	99.7%	Score 660;	DB 10;	Length 1071;
	Best Local Similarity	100.0%;	Pred. No. 9.6e-162;		
	Matches 660;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3	CGCGCAGCAGACCCCTCTCTGCGACGCGACGCCGCCCGCACCCACCATGCGCACAGTTCA	62		
Db	42	CGCGCAGCAGACCCCTCTCTGCGACGCGACGCCGCCCGCACCCACCATGCGCACAGTTCA	101		
Qy	63	GCAGCTGGAGGAGAGATGGCGCCTGTGTGACAGCAAAAGCCTTTGATGAATACATGAAGGA	122		
Db	102	GCAGCTGGAGGAGAGATGGCGCCTGTGTGACAGCAAAAGCCTTTGATGAATACATGAAGGA	161		

RESULT 12

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 98.0%; Score 649; DB 16; Length 1072;
Best Local Similarity 99.8%; Pred. No. 7.2e-159;
Matches 660; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QV 3 CGCGGACGCAGACCCCTTCTGTGACGCCGCCGCCGCCACCAT-GGCCACAGTTC 61

42	CGCGACGAGACCCCTCTCTGACGCCAGCCGCCGACCCAGCATGGGACACAGTTTC	101
62	AGCAGCTGGAAGGAGATGCGCCCTGGTGACAGCAAAAGGCTTGCATGAATACATGAAGG	121
102	AGCAGCTGGAAGGAGATGCGCCCTGGTGGAAGCAAAAGGCTTGCATGAATACATGAAGG	161
122	AGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCCAAGCCAGAGATTGTATCA	181
162	AGCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCCAAGCCAGAGATTGTATCA	221
182	TCACCTTGATGGTAAAAACCTCACATAAAACTGAGAGCACTTTGAAAAACAACACAGT	241
222	TCACCTTGATGGTAAAAACCTCACATAAAACTGAGAGCACTTTGAAAAACAACACAGT	281
242	TTTCTTGTAACCTGGGAGAGAAGTTTGAAGAAAAACACAGCTCATGGCGAAAAAATCTCAGA	301
282	TTTCTTGTAACCTGGGAGAGAAGTTTGAAGAAAAACACAGCTCATGGCGAAAAAATCTCAGA	341
302	CTGCTGCAACTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGGAAA	361
342	CTGCTGCAACTTTACAGATGGTGCATTTGGTTCAGCATCAGGAGTGGGATGGGAAGGAAA	401
362	GCACAATPACAAGAAAAATTGAAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAATG	421
402	GCACAATPACAAGAAAAATTGAAAGATGGGAATTTAGTGGTGGAGTGTGTCATGAACAATG	461
422	TCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCACTTTGGACAGG	481
462	TCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAAAATTCATCATCACTTTGGACAGG	521
482	AGTTAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTTC	541
522	AGTTAATTAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAAATCTCCATCTGTTTCTTTTC	581
542	TTTTTTTTTTTCATCTGTTCAATATATCTTTATCATAAAAATTTTACATGCAAGCTATT	601
582	TTTTTTTTTTTCATCTGTTCAATATATCTTTATCATAAAAATTTTACATGCAAGCTATT	641
602	TCAAAGTGTGTGGATTAAATTAGGATCATCCCTTTGGTTTAAATAAATGTTGTTGTGC	661
642	TCAAAGTGTGTGGATTAAATTAGGATCATCCCTTTGGTTTAAATAAATGTTGTTGTGC	701
662	T 662	
702	T 702	

```

RESULT 13
US-09-920-455-23
; Sequence 23, Application US/09920455
; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
; CURRENT APPLICATION NUMBER: US/09/920,455
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 528..553, 579..581
; OTHER INFORMATION: n = A,T,C or G
US-09-920-455-23

```

[illegible]

```

RESULT 14
US-09-920-455-147
; Sequence 147, Application US/09920455
; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
; CURRENT APPLICATION NUMBER: US/09/920,455
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 550..576
; OTHER INFORMATION: n = A,T,C or G
US-09-920-455-147

```

QY 6 CGACGCAGACCCCTCTCTGCACGCCAGCCCGCCCGCACCATGGCCACAGTTCAGCA 65

Db 22 CGACGGAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATGCGCCACAGTTCAGCA 81
Qy 66 GCTGGAAGGAGATGGCGCTGTGGACAGCAAGAGCTTTGATCAATACATGAGGAGCT 125
Db 82 GCTGGAAGGAGATGGCGCTGTGGACAGCAAGAGCTTTGATCAATACATGAGGAGCT 141
Qy 126 AGAGTGGGAATAGCTTTGGGAAATGGGCGCAATGGCCAGCCAGATTTGATCATCAC 185
Db 142 AGAGTGGGAATAGCTTTGGGAAATGGGCGCAATGGCCAGCCAGATTTGATCATCAC 201
Qy 186 TTGTGATGGTAAACCTCACATACCACTGAGAGCCTTTGAGAGCACTTTGAGAGCACTTTTTC 245
Db 202 TTGTGATGGTAAACCTCACATACCACTGAGAGCCTTTGAGAGCACTTTGAGAGCACTTTTTC 261
Qy 246 TTGTACCTCGGAGAGAGTGTGAAGAAACACAGCTGTGAGGAGAGAACTCAGACTGT 305
Db 262 TTGTACCTCGGAGAGAGTGTGAAGAAACACAGCTGTGAGGAGAGAACTCAGACTGT 321
Qy 306 CTCGCACTTTACAGATGGTGCATTTGTTGAGCATCAGGAGTGGGAGGAGAAAGCAC 365
Db 322 CTCGCACTTTACAGATGGTGCATTTGTTGAGCATCAGGAGTGGGAGGAGAAAGCAC 381
Qy 366 AATAACAGAGAAATTCGAAGATGGAAATTTAGTGGTGGAGTGTGATGAACAATGTCCAC 425
Db 382 AATAACAGAGAAATTCGAAGATGGAAATTTAGTGGTGGAGTGTGATGAACAATGTCCAC 441
Qy 426 CTGTACTCGGATCTATGAAAAAGTAGAATAAATTTCCATCATCACTTTTGGACAGAGTT 485
Db 442 CTGTACTCGGATCTATGAAAAAGTAGAATAAATTTCCATCATCACTTTTGGACAGAGTT 501
Qy 486 AATAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAACTCCATCATCACTTTTCTTTT 545
Db 502 AATAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAACTCCATCATCACTTTTCTTTT 561
Qy 546 TTTTTCATCTACTGTCTCAATTTCTTTTATCAATCAATTTTAC 590
Db 562 TTTTTCATCTACTGTCTCAATTTCTTTTATCAATCAATTTTAC 606

RESULT 15

US-09-981-151A-17
; Sequence 17, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Beha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17

; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 634
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-981-151A-17

Query Match 80.9%; Score 535.6; DB 10; Length 634;
Best Local Similarity 92.6%; Pred. No. 2.3e-129;
Matches 589; Conservative 0; Mismatches 34; Indels 13; Gaps 2;

Qy 3 CGCGCAGCAGACCCCTCTCTGACGCGCAGCCCGCCGACCCACCATGCGCCACAGTTCAC 62
Db 12 CGCGCAGCAGACCCCTCTCTGATGCGCGCGCCGCGCGCCACCATGCGCCACAGTTCAC 71
Qy 63 GCAGCTGGAAGGAGATGGCGCTGTGACAGCAGCAAGCTTTGATCAATACATGAGGA 122
Db 72 GCAGCTGGAAGGAGATGGCGCTGTGATGAGCAAGCTTTGATCAATACATGAGGA 131
Qy 123 GCTAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCCAAGCCAGATTTGATCAT 182
Db 132 GGGAGGAGTGGGAATAGCTTTGCGAAAAATGGCGCAATGGCCCAAGCCAGATTTGATCAT 191
Qy 183 CACTTGTGATGTTAAACCTCACATATAAACTGAGAGACCTTTGAAACAAACACAGTT 242
Db 192 CACTTGTGATGTTAAACCTCACATATAAACTGAGAGACCTTTGAAACAAACAGTT 248
Qy 243 TTCTGTACCTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAACTCAGAC 302
Db 249 TTCTGTACCTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAACTCAGAC 308
Qy 303 TGTCTGCAACTTTTACAGATGGTGCATTTGCTCAGCATCAGGAGTGGGAGGAGAAAG 362
Db 309 TGTGTGACCTTTGCGAGATGGTGCATTTGCTCAGCATCAGGAGTGGGAGGAGAAAG 368
Qy 363 CACAATAACAGAAAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTGATGAACAATGT 422
Db 369 CACAATAACAGAAAAATTTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTGATGAACAATGT 428
Qy 423 CACTGTACTCGGATCTATGAAAAAGTAGAATAAATTTCCATCATCACTTTTGGACAGGA 482
Db 429 CGCCTGTACTCGGATCTATGAAAAAGTAGAATAAATTTCCATCATCACTTTTGGACAGGA 488
Qy 483 GTTAATAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAACTTCCATCTCTTTCTTCT 542
Db 489 GTTAATAAGAGAAATGACCAAGCTCAGTTCAATGAGCAAACTTCCATCTCTTTCTTCT 539
Qy 543 TTTTTCATCTACTGTGTTCAATTTATCTTTTATCAATAACATTTTACATGACGATTTT 602
Db 540 -TTTTTTTCATCTACTGTGTTCAATTTATCTTTTATCAATAACGTTTTCATGACGATTTT 598
Qy 603 CAAAGTGTGTTGAGTAATAGGATCATCCCTTTGG 638

Db 599 CAAAGTGTCTTGGATTAAATTAGGATCATCCCTTTGG 634

Search completed: July 12, 2005, 19:05:22
Job time : 1170.11 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 15:11:38 ; Search time 130.285 Seconds
(without alignments)
8314.170 Million cell updates/sec

Title: US-09-788-074-4
Perfect score: 662
Sequence: 1 accgcgcagcagaccctc.....taataataatgtttgtgtgt 662

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	645	97.4	45762	4	US-09-949-016-16651 Sequence 16651, A
2	408	61.6	408	1	US-08-446-600A-3 Sequence 3, Appli
3	350	52.9	354	3	US-09-843-597-143 Sequence 143, App
4	350	52.9	354	4	US-09-480-884A-143 Sequence 143, App
5	350	52.9	354	4	US-09-542-615A-143 Sequence 143, App
6	350	52.9	354	4	US-09-606-421B-143 Sequence 143, App
7	350	52.9	354	4	US-09-221-107-143 Sequence 143, App
8	350	52.9	354	4	US-09-466-396A-143 Sequence 143, App
9	350	52.9	354	4	US-09-476-496A-143 Sequence 143, App
10	350	52.9	354	4	US-09-630-940B-143 Sequence 143, App
11	350	52.9	354	4	US-09-285-479-143 Sequence 143, App
12	155.8	23.5	620	4	US-09-949-016-1540 Sequence 1540, Ap
13	149.6	22.6	2150	4	US-09-949-016-389 Sequence 389, App
14	148	22.4	2149	4	US-09-949-016-1539 Sequence 1539, Ap
15	143.4	21.7	463	4	US-09-513-999C-14348 Sequence 14348, A
16	137.6	20.8	429	4	US-09-799-451-387 Sequence 387, App
17	128.2	19.4	267	4	US-09-513-999C-893 Sequence 893, App
18	120.4	18.2	731	3	US-09-043-646-1 Sequence 1, Appli
19	120.4	18.2	731	4	US-09-971-187-1 Sequence 1, Appli
20	120.4	18.2	1022	4	US-09-949-016-1203 Sequence 1203, Ap
21	115	17.4	606	4	US-09-513-999C-3961 Sequence 3961, Ap
22	113.2	17.1	493	4	US-09-513-999C-3851 Sequence 3851, Ap
23	84.4	12.7	611	4	US-09-621-976-12879 Sequence 12879, A
24	74	11.2	8524	4	US-09-949-016-13282 Sequence 13282, A
25	73.6	11.1	9739	4	US-09-949-016-12131 Sequence 12131, A
26	73.6	11.1	9739	4	US-09-949-016-13281 Sequence 13281, A
27	63	9.5	286	4	US-09-513-999C-1409 Sequence 1409, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-16651/c	7218	1	US-08-232-463-14	Sequence 14, Appl	
Sequence 16651, Application US/09949016	601	4	US-09-949-016-42044	Sequence 42044, A	
Patent No. 6812339	9.2	8597	4	US-09-949-016-12945	Sequence 12945, A
GENERAL INFORMATION:	9.2	8597	4	US-09-471-276-676	Sequence 676, App
APPLICANT: VENTER, J. Craig et al.	8.4	276	4	US-09-949-016-5760	Sequence 5760, Ap
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED	7.4	405	4	US-09-774-528-249	Sequence 249, App
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF	7.4	1287	4	US-09-513-999C-1607	Sequence 1607, Ap
FILE REFERENCE: CL001307	7.1	354	4	US-09-621-976-17115	Sequence 17115, A
CURRENT FILING DATE: 2000-04-14	7.1	458	4	US-08-847-724-2	Sequence 2, Appli
PRIOR APPLICATION NUMBER: 60/241,755	6.6	466	4	US-09-702-705-61	Sequence 61, Appl
PRIOR FILING DATE: 2000-10-20	6.6	466	4	US-09-736-457-61	Sequence 61, Appl
PRIOR APPLICATION NUMBER: 60/237,768	6.6	466	4	US-09-614-124B-61	Sequence 61, Appl
PRIOR FILING DATE: 2000-10-03	6.6	466	4	US-09-671-325-61	Sequence 61, Appl
PRIOR APPLICATION NUMBER: 60/231,498	6.6	466	4	US-09-589-184-61	Sequence 61, Appl
PRIOR FILING DATE: 2000-09-08	6.6	924	1	US-09-658-824-61	Sequence 61, Appl
NUMBER OF SEQ ID NOS: 207012	6.6	924	2	US-08-468-709B-1	Sequence 1, Appli
SOFTWARE: FastSeq for Windows Version 4.0	6.6	924	1	US-08-241-664B-1	Sequence 1, Appli
SEQ ID NO 16651	6.6	924	4	US-09-640-173-174	Sequence 174, App
LENGTH: 45762					
TYPE: DNA					
ORGANISM: Human					
US-09-949-016-16651					
Query Match	97.4%	Score 645;	DB 4;	Length 45762;	
Best Local Similarity	100.0%;	Pred. No. 1.2e-178;			
Matches	645;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	18	CTCTCTGCACGCGCCGCCGCCACCCACACGCTTTCATGAATACATGAGGAGCTAGGAGTGGGAAT	77		
Db	28146	CTCTCTGCACGCGCGCCGCCGCCACCCACACGCTTTCATGAATACATGAGGAGCTAGGAGTGGGAAT	28087		
QY	78	ATGGCGCCCTGTGGGAGCAGCAAAAGGCTTTCATGAATACATGAGGAGCTAGGAGTGGGAAT	137		
Db	28086	ATGGCGCCCTGTGGGAGCAGCAAAAGGCTTTCATGAATACATGAGGAGCTAGGAGTGGGAAT	28027		
QY	138	AGCTTTTGCAGAAATATGGGCGCAATGGCCAAAGCCAGATTTGTATCATCTGATGATGATAA	197		
Db	28026	AGCTTTTGCAGAAATATGGGCGCAATGGCCAAAGCCAGATTTGTATCATCTGATGATGATAA	27967		
QY	198	AAACCTCACCATAAACTGAGGACACTTTGAAAACAACACAGTTTTTCTGTACCTGGG	257		
Db	27966	AAACCTCACCATAAACTGAGGACACTTTGAAAACAACACAGTTTTTCTGTACCTGGG	27907		
QY	258	AGAGAAGTTTGAAGAAACACACAGCTGATGGCGAAAAAATCTCAGACTGTCTGCAACTTTAC	317		
Db	27906	AGAGAAGTTTGAAGAAACACACAGCTGATGGCGAAAAAATCTCAGACTGTCTGCAACTTTAC	27847		
QY	318	AGATGCTGCAATTGGTTTCAGCATCAGGAGTGGGATGGGAAGGAAAGCACATAACAGAAA	377		

Db 27846 AGATGGTCATGGTTCAGCATCAGAGTGGATGGAGGAGGAGCAACAATAACAGAAA 27787
Qy 378 ATTGAAGATGGGAATATTAGTGGTGGAGTGTGCATGAACAATGTCACCTGTACTCGGAT 437
Db 27786 ATTGAAGATGGGAATATTAGTGGTGGAGTGTGCATGAACAATGTCACCTGTACTCGGAT 27727
Qy 438 CTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGAGTTAATTAGAGAA 497
Db 27726 CTATGAAAAAGTAGAATAAAAAATTCATCATCATCTTTGGACAGGAGTTAATTAGAGAA 27667
Qy 498 GACCAAGCTCAGTTCCAATGAGCAAAATCTCCATCTCTTTCTTTCTTTTTCATTAC 557
Db 27666 GACCAAGCTCAGTTCCAATGAGCAAAATCTCCATCTCTTTCTTTCTTTTTCATTAC 27607
Qy 558 TGTGTTCAATTATCTTTATCATATAAATTTTACATGCAGCTATTTCAAAGTGTGTGGAT 617
Db 27606 TGTGTTCAATTATCTTTATCATATAAATTTTACATGCAGCTATTTCAAAGTGTGTGGAT 27547
Qy 618 TAAATTAGATCATCCCTTTGGTTAATAAATAAATGTTTGTGCT 662
Db 27546 TAAATTAGATCATCCCTTTGGTTAATAAATAAATGTTTGTGCT 27502

RESULT 2
US-08-446-600A-3
; Sequence 3, Application US/08446600A
; Patent No. 5719126
; GENERAL INFORMATION:
; APPLICANT: No. 5719126dlund, James J. and Farooqui, Jamal Z.
; TITLE OF INVENTION: MELANOGENIC INHIBITOR, AND METHODS OF PRODUCING AND USING THE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; STREET: 2500 PNC Center, 201 East Fifth St.
; CITY: Cincinnati
; STATE: OH
; COUNTRY: USA
; ZIP: 45202-4182
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,600A
; FILING DATE: 24 May 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/980,513
; FILING DATE: 24 No. 5719126ember 1992
; APPLICATION NUMBER: PCT/US93/11139
; FILING DATE: 16 No. 5719126ember 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Ann G. Robinson
; REGISTRATION NUMBER: 39,820
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 651-6128
; TELEFAX: (513) 651-6981
; TELEX: 21-4396 F&J Cin
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 bases
; TYPE: nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
US-08-446-600A-3

Query Match 61.6%; Score 408; DB 1; Length 408;
Best Local Similarity 100.0%; Pred. No. 5.2e-110;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 ATGCCACAGTTTCAGCAGCTGGAAGGAGATGGGCGCTGTGGTGCAGCAGCAAAAGGCTTTGAT 108
Db 1 ATGCCACAGTTTCAGCAGCTGGAAGGAGATGGGCGCTGTGGTGCAGCAGCAAAAGGCTTTGAT 60
Qy 109 GAATACATGAAGGAGCTAGGAGTGGGAATAGCTTTTGGAAAAATGGGCGCAATGGCCAAAG 168
Db 61 GAATACATGAAGGAGCTAGGAGTGGGAATAGCTTTTGGAAAAATGGGCGCAATGGCCAAAG 120
Qy 169 CCAGATTGTATCATCACTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTG 228
Db 121 CCAGATTGTATCATCACTTGTGATGGTAAAAACCTCACCATAAAAACTGAGAGCACTTTG 180
Qy 229 AAAACAACACAGTTTTCTTTGACCTGGGAGAGAAATTTGAAGAAACCAACAGCTGATGGC 288
Db 181 AAAACAACACAGTTTTCTTTGACCTGGGAGAGAAATTTGAAGAAACCAACAGCTGATGGC 240
Qy 289 AGAAAACTCAGACTGTCTGCACTTTTACAGATGGTGTGATTTTTCAGCATCAGAGTGG 348
Db 241 AGAAAACTCAGACTGTCTGCACTTTTACAGATGGTGTGATTTTTCAGCATCAGAGTGG 300
Qy 349 GATGGAGGAGGAGCAACAATAACAAGAAATTTGAAGATGGGAAATTTAGTGTGGAGTGT 408
Db 301 GATGGAGGAGGAGCAACAATAACAAGAAATTTGAAGATGGGAAATTTAGTGTGGAGTGT 360
Qy 409 GTCATGAACAATGTCACCTGTACTCGGATCTATGAAAAAGTAGAATAA 456
Db 361 GTCATGAACAATGTCACCTGTACTCGGATCTATGAAAAAGTAGAATAA 408

RESULT 3
US-09-643-597-143
; Sequence 143, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 143
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-643-597-143

Query Match 52.9%; Score 350; DB 3; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.7e-93;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 281 CTGATGGCAGAAAAAAGTCTGCTCAACTTTTACAGATGGTGTGATTTTTCAGCATC 340
Db 5 CTGATGGCAGAAAAAAGTCTGCTCAACTTTTACAGATGGTGTGATTTTTCAGCATC 64
Qy 341 AGGAGTGGGATGGGAAGGAGGAGCAATAACAAGAAAAATTTGAAGATGGGAAATTTAGTGG 400
Db 65 AGGAGTGGGATGGGAAGGAGGAGCAATAACAAGAAAAATTTGAAGATGGGAAATTTAGTGG 124
Qy 401 TGGAGTGTGTGATGAACAATGTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAT 460
Db 125 TGGAGTGTGTGATGAACAATGTCACCTGTACTCGGATCTATGAAAAAGTAGAATAAAT 184

Query Match 52.9%; Score 350; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.7e-93;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 340
DB 5 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 64

QY 341 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 400
DB 65 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 124

QY 401 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 460
DB 125 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 184

QY 461 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 520
DB 185 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 244

QY 521 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 580
DB 245 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 304

QY 581 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 630
DB 305 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 354

RESULT 10

US-09-630-940B-143
; Sequence 143, Application US/09630940B
; Patent No. 6737514

GENERAL INFORMATION:
; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Tongtong

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.

; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C10

; CURRENT APPLICATION NUMBER: US/09/630,940B

; CURRENT FILING DATE: 2000-08-02

; NUMBER OF SEQ ID NOS: 367

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 143

; LENGTH: 354

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-630-940B-143

Query Match 52.9%; Score 350; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.7e-93;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 340
DB 5 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 64

QY 341 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 400
DB 65 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 124

QY 401 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 460
DB 125 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 184

QY 461 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 520
DB 185 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 244

QY 521 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 580
DB 245 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 304

QY 581 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 630
DB 305 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 354

RESULT 11

US-09-285-479-143

; Sequence 143, Application US/09285479

; Patent No. 6821518

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF

; FILE REFERENCE: 210121.455C3

; CURRENT APPLICATION NUMBER: US/09/285,479

; CURRENT FILING DATE: 1999-04-02

; NUMBER OF SEQ ID NOS: 172

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 143

; LENGTH: 354

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-285-479-143

Query Match 52.9%; Score 350; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.7e-93;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 340
DB 5 CTGATGGCAGAAAACTCAGACTGCTGCAACTTTTACAGATGGTGCATTGGTTTCAGCATC 64

QY 341 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 400
DB 65 AGGAGTGGGATGGGAAGGAAGCACAATAACGAAGAAATTTGAAAGATGGGAAATTAGTGG 124

QY 401 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 460
DB 125 TGGAGTGTCTCATGAACATGTCACCTGTCTGCGATCTATGAAAGATAGATAAAAT 184

QY 461 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 520
DB 185 TCCATCATCACTTTGGACAGGAGTTAAATTAAGAGAATGACCAAGCTCAGTTCAATGAGCA 244

QY 521 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 580
DB 245 AATCTCCATACACTGTTCTTCTTTCTTTTTCATTACTGTGTTCAATTAATCTTTATCATA 304

QY 581 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 630
DB 305 AACATTTTACATGCAGCTATTTCAAAGTGTGTTGGATTAATTAGGATCAT 354

RESULT 12

US-09-949-016-1540

; Sequence 1540, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1540
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1540

Query Match 23.5%; Score 155.8; DB 4; Length 620;

Best Local Similarity 62.4%; Pred. No. 1.2e-35;
Matches 244; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

```
QY 68 TGAAGGAGGATGGCGCTGGTGGACAGCAAGGCTTTTGATGAATACATGAAGGAGCTAG 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TTGTAGGTACCTGGAACTTGCTCCAGTGAAACTTTTGATGATTAATGAAGAGTAG 120

QY 128 GAGTGGGAATAGCTTTGGGAAAATGGGCGCAATGGCCCAAGCCAGATTGTATCATCTT 187
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GAGTGGGCTTTGCCACAGGAAGTGGCTGGCATGGCCAAACCTAAATGATCATCAGTG 180

QY 188 GTGATGGTAAACCTCCACATAAACTGAGAGCACATTTGAAACAACACAGTTTCTT 247
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TGAATGGGGGTGTGATCACTCACTAAATCTGAAAGTACCTTTTAAATAATCTGAGATTTCCT 240

QY 248 GTACCTCGGAGAGAAGTTTGAAGAAACACACAGCTGATGGCGAGAAAACCTCAGACTGTCT 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TCATCTAGTGGCCAGGAATTTGACGAAGTCACTGAGATGACAGGAAGTCAAGGACCA 300

QY 308 GCACTTTACAGATGGTGATTTGGTTCAGCATCAGAGTGGGATGGGAGGAAAGCAAA 367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TAACTTTAGTGGGGTGTCTGGTACATGTGCAGAAATGGGATGGAAATCAACCA 360

QY 368 TAAAGAAAATGAAGATGGGAATTTAGTGGTGGAGTGTCTCATCAATGTCACT 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TAAAGAGAAAACGAGAGGATGATAAATCTGGTGGTGGATGGCTCATGAAGCGTCACTT 420

QY 428 GTACTCGGATCTATCAAAAAGTAGATAAAA 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 CCACAGAGTTTATGAGAGCATTAAGCCAA 451
```

RESULT 13
US-09-949-016-389
; Sequence 389, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 389
; LENGTH: 2150
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-389

Query Match 22.6%; Score 149.6; DB 4; Length 2150;

Best Local Similarity 54.6%; Pred. No. 1.6e-33;
Matches 299; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

```
QY 73 GGAAGATGGCCCTGGTGGACAGCAAGGCTTTGATGAATACATCAGGAGCTAGGACTG 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 GGCACCTGGAACTTGCTCTGACGAGAACTTTTACGATTAACAGAACTCTCTTCAAG 112

QY 133 GGAATAGCTTTGGGAAAATGGGCGCAATGGCCCAAGCCAGATTGTATCATCACTTTGTAT 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 GGGTTAGCCACAGAAAATTTGGCCAAACCCACTGTGATCATCAGCAAGAA 172

QY 193 GGTAAAAACCTCACATAAAAACTGAGAGCACTTTGAAAAACAACACAGATTTCCTGTACC 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 GGAGATATTATACTATACGAACCTGAAAGTACCTTTTAAAAATACAGAAATCTCTTCAAG 232

QY 253 CTGGGAGAGAGTTTGAAGAAACACACAGCTGATGGCAGAAAACTCAGACTGTCTGCAAC 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 CTAGGCCAGGAATTTGAAGAAACACACAGCTGACCAATAGAAAAGACCAAGCATCTGAACC 292

QY 313 TTTACAGATGTGATTGGTTTTCAGCATCAGAGTGGGATGGGAGGAAAGCAATAACA 372
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 CTGACAGAGAGTCACTGAATCAAGTGCAGAGATGGGATGGCAAGACCAACATAAG 352

QY 373 AGAAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTGATGAACAATGTCACTGTACT 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 AGAAAGCTAGTGAATGGGAAATTTAGTGGGAAATTTGAAAGTGAAGGGCGTGTGTGCAAC 412

QY 433 CGGATCTATGAAAAGTAGATAAAAATTCATCATCATCTTTGGACAGAGGTTAATTAAG 492
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 AGAATCTATGAGAGGCTGAAAATTCATTTCTCATTTGAAGTGGCTTTTATCATTTAA 472

QY 493 AGAATGACCAAGCTCAGTTCAATGAGCAAACTCTCATACTGTCTTTCTTTTCTTTTTC 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 TGATGGAATCAATGCTTCCATTCACAAAATGATGATCACTGCAAAATTTGTTTTCG 532

QY 553 ATTACTGTGTTCAATTAATCTTTTATCATAAACATTTTACATGAGCTATTTCAAAGTGTGT 612
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 TTTTGTCTTAATATATCATGATATCAAGAGGCTTAAACTGAGAAATTAATCTAAAGTCAGT 592

QY 613 TGGATTA 620
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 GTTATTTA 600
```

RESULT 14
US-09-949-016-1539
; Sequence 1539, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1539
; LENGTH: 2149
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1539

Query Match 22.4%; Score 148; DB 4; Length 2149;

Best Local Similarity 54.4%; Pred. No. 4.6e-33;
Matches 298; Conservative 0; Mismatches 250; Indels 0; Gaps 0;

QY 73 GGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGATATACATGAAGAGCTAGAGTG 132
Db 53 GGCACCTGGAACCTTCTCTAGTGAGAACTTTGACGATTTACATGAAGCTCTGGGTGTG 112
QY 133 GGAATAGCTTTGGGAAAAATGGCGCAATGGCCAGCCAGATTGTATCATCACTTGTGAT 192
Db 113 GGGTTAGCCACCGAAAACTGGGAATTTGGCCAAACCCACTGTGATCATCAGCAAAA 172
QY 193 GGTAAAAACCTCACCAATAAACTGAGAGCACTTTTGAAAAACAACAAGTTTCTGTACC 252
Db 173 GGGATATTATTAACATACGAACCTGAAAGTACCTTTAAAAATAACAGAAATCTCCTTCAAG 232
QY 253 CTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAAAATCTCAGACTGTCTGCAAC 312
Db 233 CTAGGCCAGGAATTTGAAGAAACCAAGCTGCAATAGAAAGCAAGAGCATCGTAACC 292
QY 313 TTATACAGATGGTCACTTGGTTAGCATCAGGAGTGGGATGGGAAGCAAAAGCAATAACA 372
Db 293 CTGCAGAGAGGATCACTGAATCAAGTCAGAGATGGGATGGCNAAGAGCAACCATTAAG 352
QY 373 AGAAAAATTGAAAGATGGGAAATTTAGTGGTGGAGTGTGTCATGAACAAATGTCACCTGTACT 432
Db 353 AGAAGCTAGTGAATGGGAAATTTAGTGGGATGTAATGAAGGCGTGGTGTGCACC 412
QY 433 CGGATCTATGAAAGATAGATAAAATTTCCATCATCACTTTGGCAGGAGTTAATTAAG 492
Db 413 AGAATCTATGAGAAGGCTCGAAAAATCAATTTCTTCAATGAAGTGGCTTTTATCATTTAA 472
QY 493 AGAATCAACCAAGCTCAGTTCAATGAGCAAAATCCATACCTGTTCTTTCTTTTCTTTC 552
Db 473 TGATGGAATCAATTCCTTCCATGTGACAAACTGAATACATCTGCAAAATATTTGTTTTGC 532
QY 553 ATTACTGTGTTCAATATCTTTATCATATAAACAATTTTACATGCAGCTATTTCAAAGTGTG 612
Db 533 TTTTGTCTTAATATATACAGATATGCAAGGCTTAACCTGAGAAATTAATCTAAAGTCACT 592
QY 613 TGAATTAA 620
Db 593 GTTATTTA 600

RESULT 15
US-09-513-999C-14348
; Sequence 14348, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14348
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 289
; OTHER INFORMATION: y=c or t
US-09-513-999C-14348

Query Match 21.7%; Score 143.4; DB 4; Length 463;
Best Local Similarity 62.4%; Pred. No. 4.4e-32;
Matches 239; Conservative 1; Mismatches 142; Indels 1; Gaps 1;
QY 68 TGGAGGAAGATGGCGCTGTGGACAGCAAAAGGCTTTGATGAATACATGAAGAGCTAG 127

Db 82 TTGTAGGTACCTGGAAACCTTGTCTCCAGTGAACCTTTGATGATTATATGAAGAAGTAG 141
QY 128 GAGTGGGAATAGCTTTTGGAAAAATGGGCGCAATGGCCAAAGCCAGATTTGTATCATCACTT 187
Db 142 GAGTGGGCTTTTGGCCACCAGGAAGTGGCTGGCATGGCCAAACCTTAACATGATCATCAGTG 201
QY 188 GTGATGTTAAAAACCTCACCAATAAACTGAGAGCACTTTGAAAAACAACAAGTTTCTT 247
Db 202 TGAATGGGGATGTGATCACCATTAATCTGAAAGTACCTTTAAAAAATACTGAGATTTCT 261
QY 248 GTACCTGGGAGAGAGTTTGAAGAAACCAAGCTGATGGCAGAAAAAATCTCAGACTGTCT 307
Db 262 TCATACTGGGCCAGGAATTTGACG-AAGYACTGCAGATGACAGGAAAGTCAAGAGCACCA 320
QY 308 GCAACTTTACAGATGGTGCATTTGGTTTCAGCATCAGGATGGGATGGGAAGGAAAGCACAA 367
Db 321 TAACTTAGATGGGCTGTCTTGTACATGTGCAGAAATGGGATGGGAATCAACCA 380
QY 368 TAAAGAAAAATTTGAAGATGGGAAATTTAGTGGTGGAGTGTGTCATGAACAATGTCACT 427
Db 381 TAAAGAGAAAAACGAGAGGATGATAAACTGGTGGTGAATGCGTTCATGAAAGCGTCACTT 440
QY 428 GTACTCGGATCTATGAAAAGTA 450
Db 441 CCACGAGAGTTTATGAGAGCA 463

Search completed: July 12, 2005, 18:08:08
Job time : 132.285 secs

This Page Blank (uspto)